

# Organization and function of the 3D genome

Boyan Bonev and Giacomo Cavalli

**Abstract** | Understanding how chromatin is organized within the nucleus and how this 3D architecture influences gene regulation, cell fate decisions and evolution are major questions in cell biology. Despite spectacular progress in this field, we still know remarkably little about the mechanisms underlying chromatin structure and how it can be established, reset and maintained. In this Review, we discuss the insights into chromatin architecture that have been gained through recent technological developments in quantitative biology, genomics and cell and molecular biology approaches and explain how these new concepts have been used to address important biological questions in development and disease.

## Xist region

Region on the X chromosome, which contains the long non-coding RNA Xist and is essential for X chromosome inactivation in placental mammals.

**Carbon copy chromosome conformation capture (5C).** Combines a proximity ligation chromosome conformation capture (3C) approach with amplification of interactions involving preselected sets of regions (typically two sets of hundreds to thousands of restriction fragments) to improve resolution.

*Institute of Human Genetics,  
UPR1142 National Centre for  
Scientific Research (CNRS);  
and University of Montpellier,  
141 Rue de la Cardonille,  
34396 Montpellier Cedex 5,  
France.*

*Correspondence to G.C.  
giacomo.cavalli@igh.cnrs.fr*  
doi:10.1038/nrg.2016.112  
Published online 14 Oct 2016;  
corrected online 31 Oct 2016

Over the past two decades, folding of DNA into chromatin has increasingly been recognized as important<sup>1–3</sup>, with several studies highlighting the significance of spatial gene positioning for essential biological functions such as transcription, replication, DNA repair and chromosome translocation<sup>4,5</sup>. However, how chromatin is folded within the nucleus is still a matter of considerable debate. At the most basic level, folding of DNA into nucleosomes is well described<sup>6</sup>, but it is still unclear how individual nucleosomes interact with one another. At the kilobase-to-megabase scale, chromatin interactions that might involve loop formation between regulatory elements are crucial for correct cell identity, but how these interactions are established and regulated is not well understood. A major recent discovery was that, beyond individual loops, chromatin is organized in distinct structural domains, which may represent functional units of the genome<sup>7–9</sup>.

Although three-dimensional (3D) architecture must be robust, it also needs to be flexible enough to allow marked changes to occur, such as those leading to mitosis. Recent results suggest that the global structural landscape remains robust to perturbation during development, but individual genes often switch between active and inactive chromosome compartments, and specific interactions both within and between chromatin domains frequently change<sup>10</sup>. With the recent publication of very high resolution genome-wide chromatin interaction maps<sup>11,12</sup> it is becoming apparent that chromatin organization is more complex than previously anticipated, and important features for development, such as enhancer-promoter interactions, subdomain organization and weak long-range interactions, can only be reliably discovered with high sequencing depth or novel techniques. Therefore,

chromatin architecture is best studied using a combination of approaches, neither of which is comprehensive on its own. Microscopy-based methods (BOX 1) are powerful and provide important information about the relative and radial positioning of genomic regions, as well as the variability of spatial DNA organization within cell populations, but these methods are usually limited to a few regions of interest. By contrast, chromosome conformation capture (3C)-based approaches (FIG 1) are genome-wide, but their results may represent a superimposition of individual genome conformations rather than one stable structure.

An alternative approach aims to reconstruct experimental Hi-C (a high-throughput derivative of 3C) maps by modelling a composite of structures (often on the basis of polymer-based simulations of chromatin). Such methods were initially used to describe the polymer state of chromatin on the basis of Hi-C data<sup>13</sup>. Modelling approaches were later used to investigate the intercellular variability of chromatin contacts within the Xist region on the basis of carbon copy chromosome conformation capture (5C) data<sup>14</sup> and of X chromosome conformations using single cell Hi-C<sup>15</sup>. Furthermore, polymer modelling was used to show that metaphase chromosomes represent a series of consecutive loops compressed into arrays<sup>16</sup>, supporting earlier microscopy observations<sup>17</sup>. More recently, the formation of such loops was proposed to involve loop-extruding complexes<sup>18,19</sup> and ‘border elements’ such as CCCTC-binding factor (CTCF).

In this Review, we discuss the insights into chromosome folding and its relation to function that have been gained through recent technological developments. We examine the different levels of chromatin organization — from chromatin loops to chromosome territories

**Box 1 | Microscopy-based techniques to visualize the genome in 3D**

Historically, the position and organization of chromosomes, domains and specific loci within the nucleus have mostly been studied using fluorescent *in situ* hybridization (FISH). FISH has been mainly limited to examining a few predetermined loci in a few hundred cells. Recent advances in the development of custom oligonucleotide arrays such as Oligopaint<sup>129,130</sup> and novel super-resolution microscopy approaches such as STORM<sup>131</sup> and PALM<sup>132</sup> have enabled direct visualization of the fine-scale structures of the genome at unprecedented resolution. Recently, a high-throughput imaging approach called HIPMap was used to identify novel factors affecting the radial positioning of different types of genomic locus within the nucleus in a large-scale and unbiased manner<sup>105</sup>. Super-resolution microscopy was used to determine the structure of the chromatin fibre at single cell level with high spatial resolution, suggesting that nucleosomes are organized in groups of various sizes and that this nucleosome density is dynamic and cell-type specific<sup>22</sup>. Furthermore, STORM was used to determine the relationship between the physical volume occupied in the nucleus and the epigenetic state of chromatin domains in *Drosophila melanogaster*, which identified differences in the compaction between active, inactive and Polycomb-repressed domains<sup>39</sup>. Another application of microscopy allows labelling of individual chromatin proteins in order to track their dynamics or labelling of specific regions of DNA by expressing sequence-specific DNA-binding proteins fused with GFP derivatives. These methods provide invaluable information about the dynamics of individual chromosome domains or of generic chromatin<sup>133,134</sup>.

Despite this spectacular progress, current microscopy-based approaches are limited to a small number of genetic loci and do not allow a comprehensive analysis of nuclear architecture of the complete genome. However, future methods will probably improve this and allow us to examine dynamic changes of three-dimensional (3D) nuclear organization during differentiation at single cell level, which snapshots of population-based chromosome conformation capture (3C) data in fixed cells cannot.

— and we address the molecular mechanisms responsible for establishing and maintaining the 3D nuclear architecture. Furthermore, we discuss how such genome organization can be robust overall, but flexible enough to undergo marked changes during development and disease. Finally, we review the interplay between gene regulation and chromatin architecture and highlight some of the important questions that remain to be addressed in this rapidly developing field.

**Hierarchical folding of chromatin**

The largest chromosomes contain hundreds of millions of base pairs that fold in nucleosomes, chromatin fibres, chromosome domains, compartments and finally in chromosome territories. Therefore, chromatin folding is a multi-scale problem and all scales need to be understood, as regulatory information resides at all levels, from the histone–DNA interactions at the sub-nucleosomal scale to the chromosome–chromosome and chromosome–lamina interactions in the nuclear space. Furthermore, this multi-level architecture can be regulated and/or exploited by a variety of components such as transcription factors, architectural proteins and non-coding RNAs in order to coordinate gene expression and cell fate.

**Nucleosome–nucleosome interactions.** At the smallest scale of chromatin organization beyond the nucleosome one finds nucleosome–nucleosome interactions. For a long time, on the basis of *in vitro* electron microscopy, nucleosomes were thought to form arrays (often called the 30 nm chromatin fibres) with either solenoid or zig-zag shapes<sup>20,21</sup>. However, the biological relevance of the

**Locus control region (LCR).** Regulatory element that brings together multiple genes into an active chromatin hub and facilitates transcription in a cell-type-specific manner.

30 nm chromatin fibre has been increasingly called into question by several independent studies<sup>19,22–24</sup>. Contrary to expectations, nucleosomes seem to be more flexible<sup>19,24</sup> and are arranged in heterogeneous groups, called ‘clutches’ in a cell-type dependent manner<sup>22</sup>.

**Chromatin loops.** A key feature of vertebrate genomes is the relatively long distances along the linear genome separating *cis*-regulatory elements, such as enhancers, from their target genes. In order to elicit its effect, an enhancer is thought to be brought into close spatial proximity with its target promoter through the formation of a ‘chromatin loop’ (FIG. 2a). One well known example is the locus control region (LCR) of the  $\beta$ -globin cluster, which interacts strongly, via long-range chromatin contacts, with its target genes in erythroid cells (where the  $\beta$ -globin gene is active) but shows little or no interaction in cells from different lineages — for example, stem or neuronal cells<sup>25</sup>. These interactions have been proposed to form an active chromatin hub, in which high local concentrations of transcription factors and RNA polymerase II (RNAPII) lead to transcription.

Long-range chromatin contacts are not limited to enhancer–promoter interactions. Spatial associations between actively transcribed co-regulated genes in mice<sup>26</sup>, between Polycomb-repressed genes in *Drosophila melanogaster*<sup>27</sup> and more recently in mammalian cells<sup>28–30</sup> have also been observed. In another type of chromatin loops called ‘gene loops’ (FIG. 2a), which have been primarily identified in yeast, the transcription termination site of a gene loops back to make contact with its own promoter<sup>31</sup>. Gene loops have been suggested to reinforce the directionality of RNA synthesis from the promoter<sup>32</sup>. A recent study using very high resolution Hi-C supports the idea of correlation between chromatin loops and transcription by showing that the anchors of cell-type-specific loops are often the promoters of differentially expressed genes and that they contain binding sites for the architectural protein CTCF<sup>11</sup>.

**Topologically associating domains.** One of the most interesting recent discoveries in this field was that chromosomes are spatially segregated into sub-megabase scale domains, often called topologically associating domains (TADs)<sup>7–9</sup>. TADs typically manifest as contiguous square domains along the diagonal of Hi-C maps (or triangles as represented in FIG. 2b), in which regions within the same TAD interact with each other much more frequently than with regions located in adjacent domains (FIG. 2b).

The spatial partitioning of the genome into TADs correlates with many linear genomic features such as histone modifications, coordinated gene expression, association with the lamina and DNA replication timing<sup>9</sup>. Furthermore, enhancer–promoter interactions seem to be mostly constrained within a TAD<sup>33</sup>. Whereas initially mammalian TADs were identified with a median size of ~880 kb<sup>9</sup>, subsequent analysis of higher resolution Hi-C data<sup>11</sup> suggested a smaller median domain size of ~185 kb (range 40 kb–3 Mb). Strikingly, these smaller mammalian domains resemble TADs identified in

*D. melanogaster*<sup>7</sup>, both in size and in that most domains can be associated with a prominent epigenomic signature; for example, active chromatin, chromatin repressed by Polycomb, heterochromatin or association with the nuclear lamina. In addition, *D. melanogaster* TADs were proposed to correspond to the bands in polytene chromosomes<sup>34</sup>, connecting Hi-C defined regions to previous microscopy observations. In mammals, strong chromatin loops are observed at the borders of ~39% of the domains, leading to the term 'loop-domain' (REF. 11). The latter observation suggests a strong relationship between chromatin loop formation and the demarcation of domain boundaries. What distinguishes 'background' contacts, such as those among random points within a TAD, from regulatory or structural chromatin loops may be the stability of the loop, which might be increased by the binding of specific factors promoting loop formation.

TAD boundaries are enriched for insulator proteins such as CTCF (detected at ~76% of all boundaries), active transcription marks such as H3K4me3 and H3K36me3, nascent transcripts, housekeeping genes (present in ~34% of TAD boundaries), and repeat elements<sup>9</sup>. However, at least for the Xist locus, TAD organization does not seem to be a consequence of chromatin marks and was unchanged in *G9a* (also known as *EHMT2*)<sup>-/-</sup> and embryonic ectoderm development (*Eed*)<sup>-/-</sup> mutant mouse embryonic stem (ES) cells, which lack the H3K9me2 and H3K27me3 marks, respectively<sup>8</sup>. In *D. melanogaster*, transcription seems to be a better predictor of TAD boundaries than CTCF, which suggests that different organisms may have different strategies to specify chromatin domains<sup>35</sup>.

TADs are thought to be conserved between different cell types<sup>9,11</sup> and across species; however, the extent of this conservation is unclear. Much of the uncertainty seems to arise from the nested structure of mammalian TADs, whereby large TADs can be further subdivided into smaller domains (sometimes called subTADs)<sup>36,37</sup>. As a result of this hierarchical organization, how domains are identified and classified depends strongly on the resolution of the Hi-C experiment and to some extent on the method used. Importantly, and partly because of these reasons, different authors have used different nomenclature for chromosomal domains in the megabase or sub-megabase size range and a unifying definition will be hard to reach. Earlier studies using lower resolution Hi-C found 2,200 domains, of which 50–72% are conserved in different cell types and 54–76% are conserved between mouse and human cells<sup>9</sup>. Using higher resolution Hi-C, 9,274 domains were reported in GM12878 cells, of which 54% are conserved in other cell types (the evolutionary conservation of smaller domains was not reported)<sup>11</sup>. However the conservation rate might be underestimated as only a small proportion of the cell-type-specific boundaries showed clear differences in insulation (inhibition of inter-domain contacts) between different cell types<sup>9</sup>. It will be important to investigate what features demarcate these dynamic boundaries compared with the majority of stable elements in order to understand boundary formation.

#### Insulator proteins

Often present at, but not limited to, domain boundaries, insulator proteins are thought to block the interactions between regulatory elements such as enhancers and promoters. In mammals the main insulator protein is CCCTC-binding factor (CTCF), whereas in *Drosophila melanogaster* at least five different classes of insulator are known.

**Compartmentalization of megabase-scale chromatin.** At least in mammals, long-range interactions between TADs that can be located at variable distances, sometimes very far on the linear genome, give rise to compartments (FIG. 2c). Initially two types of compartment, called A and B, were identified on the basis of their preferential interaction with each other (domains in compartment A interact mostly with other type A domains, and vice versa)<sup>13</sup>. Recently, higher resolution Hi-C suggested that these two major compartments can be further subdivided into six different subcompartments (two for the active A compartment and four for the inactive B compartment)<sup>11</sup> — an observation that was further confirmed by chromosome conformation capture-on-chip (4C) experiments<sup>37</sup> and, more recently, by extensive DNA fluorescent *in situ* hybridization (FISH) studies<sup>38</sup>. Importantly, whereas TADs are mainly conserved between different cell types, compartments are not and TADs can switch between compartments A and B in a cell-type-specific manner<sup>10,13</sup>. Although it is clear that multiple TADs form a compartment, what drives this process and what is the functional distinction between a TAD and a subcompartment is less understood. It is tempting to speculate that local mechanisms such as CTCF binding and gene expression underlie TAD formation, whereas subcompartments are formed by attraction and/or repulsion between individual TADs with similar epigenetic marks. This model is supported by the strong correlation between chromatin marks of loci within a TAD compared with across TADs<sup>7,11</sup> and the observation that many TAD boundaries also demarcate subcompartment transitions<sup>11</sup>. In addition, super-resolution microscopy uncovered remarkable differences in the spatial interactions between neighbouring TADs with different epigenetic states<sup>38</sup>, showing in particular that Polycomb-repressed domains are particularly condensed and exclude neighbouring domains to a large extent<sup>39</sup>. Further genome-wide experiments in mutants deficient in Polycomb and other chromatin modifiers are required to determine the role of epigenetic marks in genome architecture.

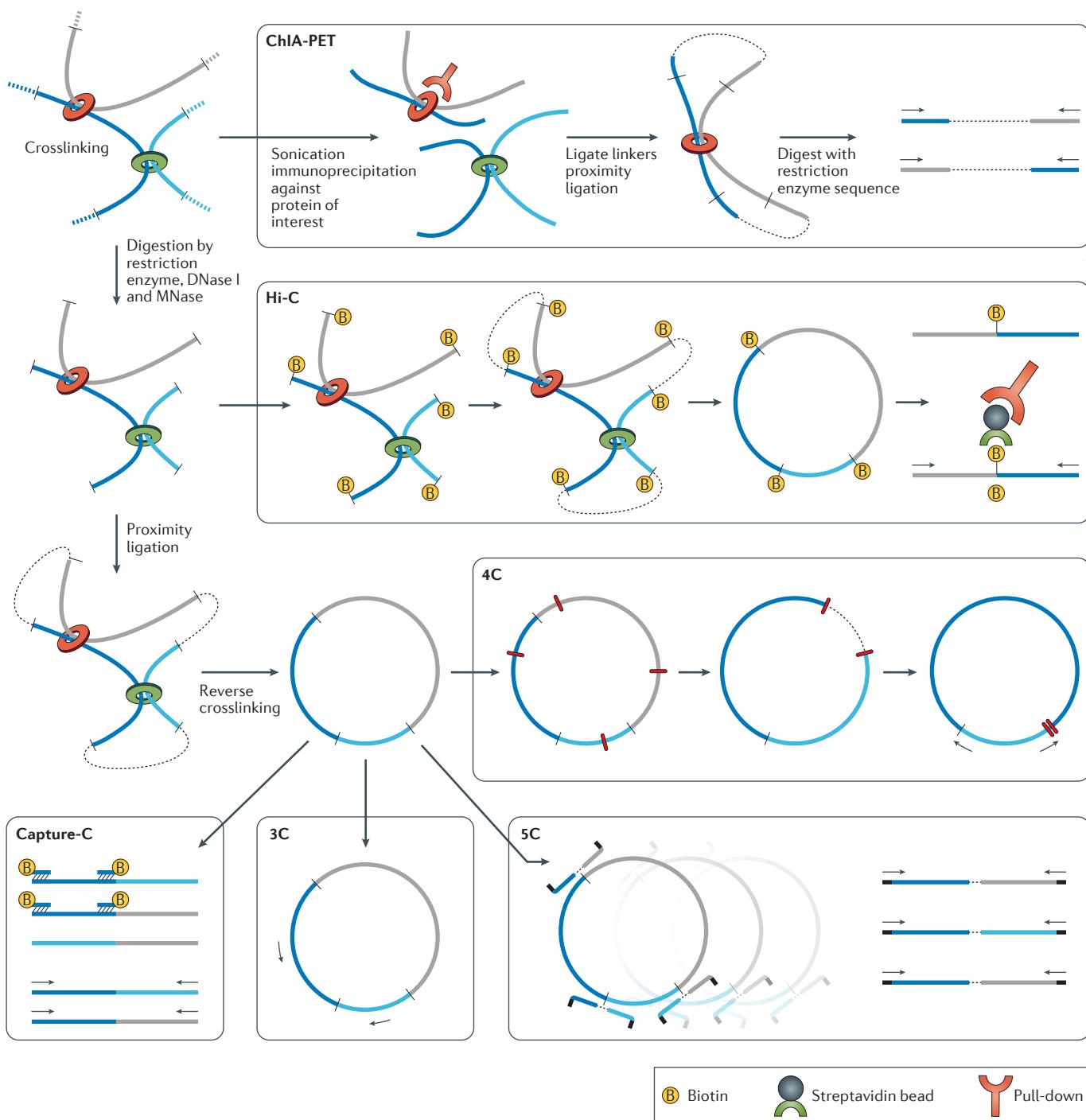
At even larger scales, chromatin is organized into individual chromosome territories (one for each chromosome), which rarely intermix (FIG. 2d). This observation, initially coming from FISH studies<sup>40,41</sup>, was later validated by genome-wide Hi-C data, which showed that interactions between loci on the same chromosome are much more frequent than contacts in *trans* between different chromosomes<sup>13</sup>.

All of these data can be summarized to conclude that chromosome architecture is formed in a hierarchical manner. First, dynamic nucleosome contacts form clutches and fibres. These engage in dynamic longer distance loops. Some of these loops that are established or stabilized by protein–protein contacts involving architectural (that is, CTCF and cohesin) and/or regulatory components (that is, transcription factors, Polycomb and heterochromatin proteins) give rise to structural landmarks, such as gene domains and TADs. Interaction among TADs of the same epigenomic type forms compartments and coalescence of compartments in the same chromosome forms chromosome territories.

### Mechanisms to organize chromatin in 3D

An important question in chromatin biology is how the structural features of 3D chromatin organization are established, maintained and potentially reset during cell cycle, development and signalling. Different species seem to deploy different components in order to establish chromosome domains. In flies, several architectural proteins are enriched at different subsets of TAD boundaries<sup>42,43</sup>, allowing dynamic regulation of each of these subsets of TAD boundaries to occur independently. In vertebrates, a partially different set of factors may fulfil

a similar function<sup>36,44</sup>. Furthermore, these proteins may establish chromosomal domains in addition to their role in other biological processes such as cell cycle and transcription<sup>45–47</sup>. Recent analysis of 76 DNA-binding proteins identified subunits of the cohesin complex, CTCF, yin yang 1 (YY1) and zinc finger protein 143 (ZNF143) as highly enriched at the anchors of strong chromatin interactions<sup>11</sup>. Together with the mediator complex, which has a well-known role in bridging enhancers and promoters in the 3D space<sup>48</sup>, both CTCF<sup>49</sup> and cohesin<sup>50</sup> have been shown to be essential for chromatin looping



**Pre-initiation complex (PIC).** Large, multi-subunit protein complex that helps recruit RNA polymerase II (RNAPII) to transcription start sites and that is required for transcription.

**Bilaterians**  
All multicellular animals with bilateral symmetry.

and they have been proposed to function combinatorially as architectural proteins to link facultative or constitutive chromosome architecture to gene regulatory outputs<sup>36</sup>.

**Architectural proteins: mediator.** How do architectural proteins bring linearly distant loci together to form a loop? Mediator is found at both the enhancers and the promoters of actively transcribed genes<sup>48</sup> and promotes transcription by enabling pre-initiation complex (PIC) assembly and RNAPII elongation (reviewed in REF. 46). In the context of 3D chromatin architecture, it

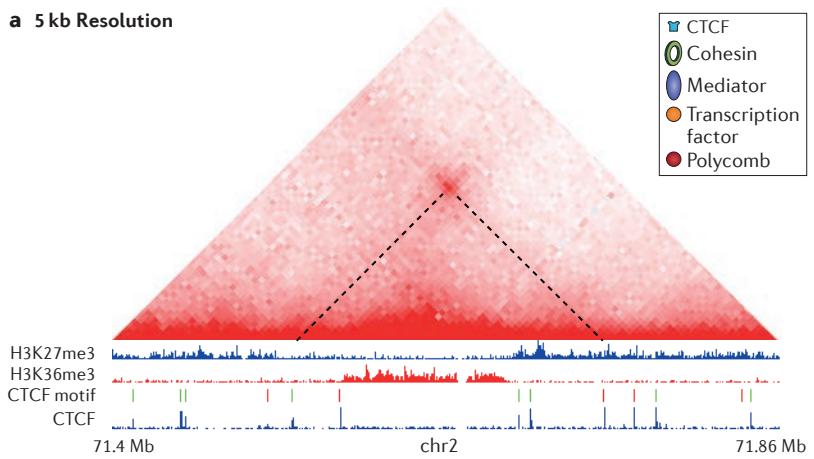
has been proposed to interact with cohesin in order to bring enhancers and promoters into physical proximity (FIG. 3a). Importantly, depletion of mediator with RNAi has been shown to diminish the strength of chromatin looping<sup>36,48,51</sup>, suggesting that it is necessary for at least a subset of interactions. As mediator is essential for transcription, it will be difficult, but important, to disentangle its involvement in looping versus RNAPII-associated transcription.

**Architectural proteins: cohesin.** Another protein with a dual functional role is cohesin. Cohesin is important for genome stability in dividing cells and is involved in sister chromatid cohesion and DNA repair (reviewed in REF. 47). In the context of chromatin architecture, cohesin interacts with both CTCF<sup>52</sup> and mediator<sup>48</sup> and is proposed to be a part of the loop-extrusion complex (discussed below) in interphase cells (FIG. 3b). Given its putative role in chromatin looping<sup>50</sup>, somewhat perplexing results were obtained in two studies examining the global chromatin architecture in cohesin-deficient post-mitotic cells; surprisingly, TADs remained mostly intact, whereas inter-domain interactions were increased and intra-domain cohesin- and CTCF-anchored loops were disrupted<sup>53,54</sup>. Importantly, in both studies, the analysed cells still contained ~10% residual cohesin, which might have been sufficient for the formation of TAD boundaries. Therefore, using systems that fully abrogate cohesin will be required to resolve its role in TAD formation.

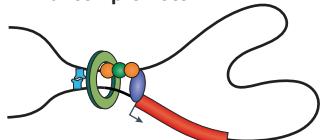
**Architectural proteins: CTCF.** In the context of architectural proteins CTCF has received perhaps the most attention recently (reviewed in depth in REF. 55). CTCF was originally characterized as an insulator protein, capable of restricting enhancer–promoter interactions both in reporter plasmids and in their native environment<sup>56,57</sup>. It is conserved in most bilaterians<sup>58</sup>, is ubiquitously expressed and is essential for embryonic development<sup>59,60</sup>. CTCF contains an 11-zinc-finger DNA-binding domain, which recognizes a specific non-palindromic motif<sup>55</sup>. In support of the role of CTCF as a barrier element, a deletion of a CTCF-binding site within the *HoxA* gene cluster affected the distribution of active compared with repressed chromatin in the two adjacent domains and resulted in the aberrant upregulation of a normally repressed gene during differentiation<sup>61</sup>. Consistent with the insulator role of CTCF, it is enriched at TAD boundaries in mammals and in *D. melanogaster*<sup>7,9</sup> (FIG. 3c). However, only 15% of all mammalian CTCF-binding sites are located within a boundary; the majority lie within TADs and are thought to be involved in intra-TAD interactions<sup>62</sup>, suggesting that CTCF binding alone may be insufficient for the establishment of boundaries. Consistent with this, CTCF knockdown in human cell lines did not strongly affect TAD boundaries but decreased intra-domain interactions and increased inter-domain contacts<sup>63</sup> (FIG. 4), although once again the data reflect an incomplete depletion of CTCF. Between ~30% and 60% of CTCF-binding sites are cell-type specific<sup>64</sup> and changes in DNA methylation at these variable sites are often correlated with differential CTCF binding.

◀ **Figure 1 | 3C-based approaches to study chromatin architecture.** Detecting DNA fragments that preferentially interact together on the basis of their proximity in the three-dimensional (3D) space was first used in 1993 (REF. 135) and subsequently improved and expanded in 2002 (REF. 136) to form the basis of all chromosome conformation capture (3C) technologies, including Hi-C (a high-throughput derivative of 3C)<sup>13</sup>. The first step of most 3C-based methods involves the formaldehyde crosslinking of cells. In most downstream protocols this is followed by fragmentation of the chromatin by digestion with a restriction enzyme, or, in a variation of 3C called chromatin interaction analysis by paired-end tag sequencing (ChIA-PET), by sonication. In ChIA-PET the next steps involve enrichment for interactions mediated by a protein of interest by immunoprecipitation, ligation of adaptors to the restriction fragment ends followed by proximity ligation, fragmentation by restriction enzyme digestion, isolation of paired-end tags (PETs) containing adaptors and paired-end sequencing<sup>62,71,137</sup>. In standard 3C-based protocols the digestion by restriction enzymes such as HindIII or DpnII is then followed by proximity-based ligation of adjacent DNA ends and determination of pair-wise interactions using either PCR or sequencing approaches. Different chromatin fragmentation methods (for example, digestion with DNase I) were recently used to improve resolution and to reduce the potential biases of standard 3C techniques<sup>138</sup>. After reverse crosslinking, different approaches can be used to identify the chromatin interactions. In the classical 3C method a pair of interacting loci are interrogated using quantitative PCR (qPCR) one at a time; in the chromosome conformation capture-on-chip (4C) protocol a second round of digestion and ligation is used to increase resolution, followed by inverse PCR with locus-specific primers to detect genome-wide interactions involving the locus of interest<sup>139</sup>. In the carbon copy chromosome conformation capture (5C) approach, primer sequences overlapping restriction fragment ends are ligated only when the two ends are immediately adjacent, then products are amplified and sequenced<sup>140</sup>. In Capture-C methodology, enrichment for interacting pairs is accomplished using biotin-labelled probes complementary to restriction fragment ends of interest<sup>141,142</sup>. In the Hi-C method the restriction fragment ends are labelled using biotin, ligated products are enriched using streptavidin pull-down after sonication and interactions are interrogated in a genome-wide all-versus-all unbiased manner. One recently developed method called micro-C, which uses MNase digestion to obtain nucleosome-based resolution of chromatin interactions in yeast<sup>23</sup>, highlights the potential of 3C-like approaches to examine chromatin interactions at the 150 bp–1 kb resolution and to interrogate nucleosome fibre folding at short ranges. However, the micro-C-based approach will be difficult to adopt in mammalian systems as that would require an order of magnitude deeper sequencing than the highest resolution (1 kb) Hi-C maps to date, with ~5 billion contacts<sup>11</sup>. This method could potentially be combined with an enrichment step, either using sequence-specific probes or with an antibody against a protein of interest to interrogate a region(s) of interest with very high resolution. Despite the ability of 3C-based approaches to interrogate chromatin-interaction features at the cell population level, it is still unclear what these features represent at a single cell level. As sequencing technologies continue to improve, one way to address this question would be to use longer than the standard 50–150 nucleotide paired reads, which would potentially allow the identification of multipartite chromatin interactions. An alternative way to address this is by determining chromatin interactions in single cells<sup>15</sup>. A major observation from single cell experiments is that only a subset of the contacts identified by population-average Hi-C are present within an individual cell — therefore, the typical maps obtained by 3C approaches probably represent a superimposition of all possible conformation states of a cell. This has important implications for the biological significance of chromatin contact maps and the 3D visual representation of chromosomes based on them. Notably, the resolution achieved in the only published single cell Hi-C study does not yet allow accessing contact frequencies in close proximity, but future improvements will probably lead to progress in this respect. 3C-based approaches are reviewed in REF. 143.

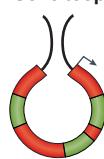
## a 5 kb Resolution



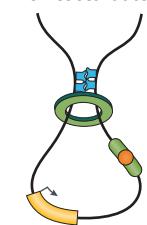
## Enhancer-promoter



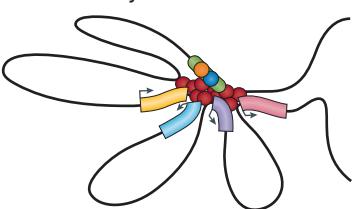
## Gene loop



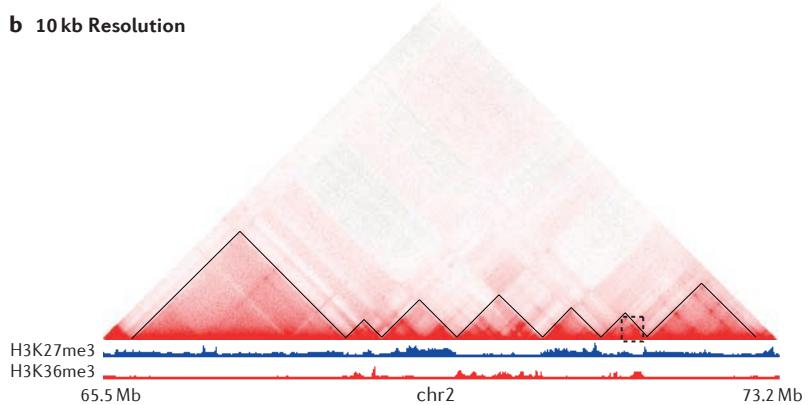
## Architectural loop



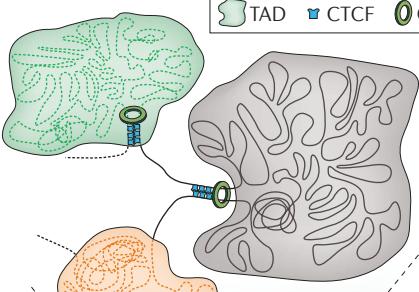
## Polycomb-mediated



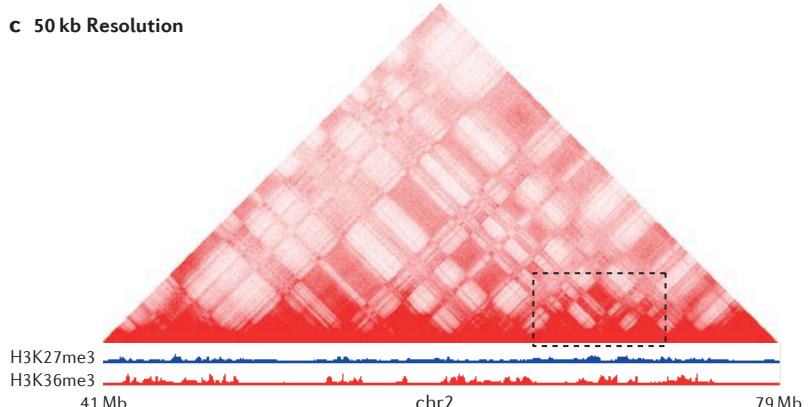
## b 10 kb Resolution



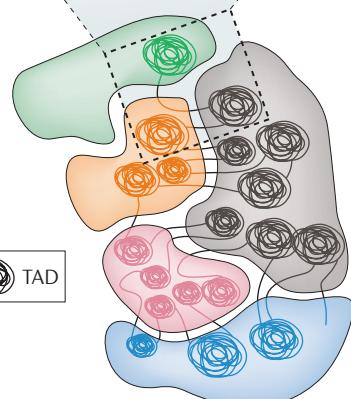
TAD CTCF Cohesin



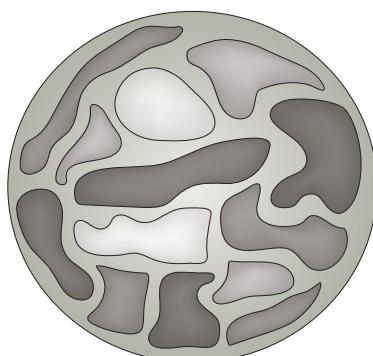
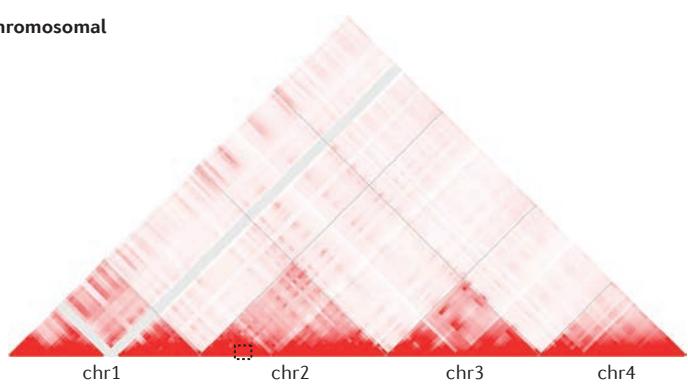
## c 50 kb Resolution



TAD



## d Interchromosomal



However, the extent to which CTCF binding is sensitive to DNA methylation and the causal relationship between these two events are still controversial. Whereas some studies<sup>64–67</sup> identified links between CTCF and methylation, CTCF binding was found to be mostly unaffected in mouse ES cells that lack DNA methylation and pre-existing DNA methylation did not block CTCF binding to a specific region<sup>68,69</sup>, suggesting that the binding of CTCF (or other transcription factors) might be causal to DNA methylation. The relationship between CTCF binding and DNA methylation thus seems to be complex and involves several feedback mechanisms. Therefore, it would be important to examine the contribution of each of these mechanisms to regulatory changes associated with 3D chromatin architecture and the consequences for gene expression during development and disease.

Another major unresolved question is whether the barrier function of CTCF (and potentially the existence of TADs) is separate from, or a consequence of the strong long-range interactions among CTCF-bound loci at boundaries of TADs. Indeed, CTCF sites at loop anchors occur predominantly in a convergent orientation, which suggests that not only binding but also directionality is important for the formation of a loop<sup>11,70</sup>. This result was confirmed by both high-resolution chromatin interaction analysis by paired-end tag sequencing (ChIA-PET)<sup>71</sup> and recent 4C analyses<sup>72</sup>. Interestingly, the directionality of transcripts in close proximity to CTCF sites was also at least partially correlated with CTCF motif orientation<sup>71</sup>, which suggests a potential role for CTCF and chromatin loops in reinforcing the directionality of RNA synthesis.

What is the significance of CTCF motif orientation for chromatin architecture? Inversion of CTCF-binding sites within the distal enhancer in the protocadherin

locus changed which promoters were targeted by the enhancer by resetting the orientation of the existing chromatin loops<sup>73</sup>. Importantly, this change in local chromatin topology was accompanied by down-regulation of the genes targeted by the endogenous loop, without a corresponding increase in the expression of the newly targeted genes. These results suggest that enhancer-anchored chromatin looping is necessary but may not be sufficient for transcription. Inversion of a CTCF motif in a loop anchor disrupted its interaction with an upstream convergent CTCF site, despite similar levels of CTCF recruitment, and, in one instance, this inversion altered the expression of the neighbouring gene<sup>72</sup>, which confirms the importance of CTCF motif orientation for looping. However, the inverted site did not engage in other chromatin loops, which suggests that loop formation may depend on the genomic context. Another study looked more globally at the consequences of CTCF motif deletion or inversion on the local domain structure<sup>19</sup>. In this case, inversion or deletion of CTCF motifs resulted in destabilization of the loop, supporting the hypothesis that convergent CTCF motif orientation is necessary for loop formation (FIG. 4). These results have important implications for the interpretation of population-based Hi-C for chromatin folding in individual cells. Specifically, they suggest that consecutive loops can and do occur simultaneously in the same cell, whereas overlapping loops (and overlapping contact domains), which are often observed in Hi-C data, probably represent alternative folding states within a cell population<sup>19</sup>. However, several important questions remain unanswered. Is CTCF binding polarity sufficient to establish a loop? Why do some convergent loops demarcate domain boundaries whereas others, located within TADs, do not? What is the contribution of the chromatin environment and transcription to loop formation?

- ◀ Figure 2 | Hierarchical organization of chromatin structure. a |** Examples of different types of chromatin loop that can potentially reside within a domain (enhancer–promoter loop, Polycomb-mediated loop, gene loop or architectural loop). On the left is an example of an architectural loop as seen in high-resolution Hi-C data (regions participating in loop formation are demarcated with dotted lines), as well as CCCTC-binding factor (CTCF)-binding profile and CTCF motif orientation (green represents forward and red represents reverse). Note that the loop is formed only between a specific forward and reverse CTCF site, despite other possible combinations. **b |** On the left is an approximately 8 Mb region containing several topologically associating domains (TADs) as seen in Hi-C maps (TADs are manually annotated with solid lines). On the right, three different TADs, enriched for either active marks (H3K4me3 and H3K36me3; grey), Polycomb (H3K27me3; green) or heterochromatin (H3K9me3; orange) are schematically represented in the three-dimensional (3D) space. CTCF proteins are shown as blue rectangles and loop-extrusion complexes (potentially cohesin) are depicted as green circles. **c |** Different topological domains with similar epigenetic signatures are characterized by stronger inter-domain interactions and are organized into compartments (blue and grey represent the active compartment, whereas interactions between green, orange and red TADs form the inactive compartment). **d |** At the highest-level of 3D organization trans-interactions are rare and individual chromosomes (chrs) occupy distinct territories (denoted by irregular shapes) within the nucleus (grey circle) — gene-rich chromosomes are preferentially found inside the nuclear core and gene-poor chromosomes are localized close to the nuclear membrane. In all panels Hi-C data are from GM12878 cells<sup>11</sup> and chromatin immunoprecipitation sequencing (ChIP-seq) tracks for H3K36me3 (red) and H3K27me3 (blue) at different resolution are shown on the left and a schematic representation of how these regions can be organized in 3D is depicted on the right. Dotted rectangles indicate the regions that were shown at higher magnification and increased resolution in the panel above. Hi-C data were visualized using the Juicebox software<sup>144</sup>.

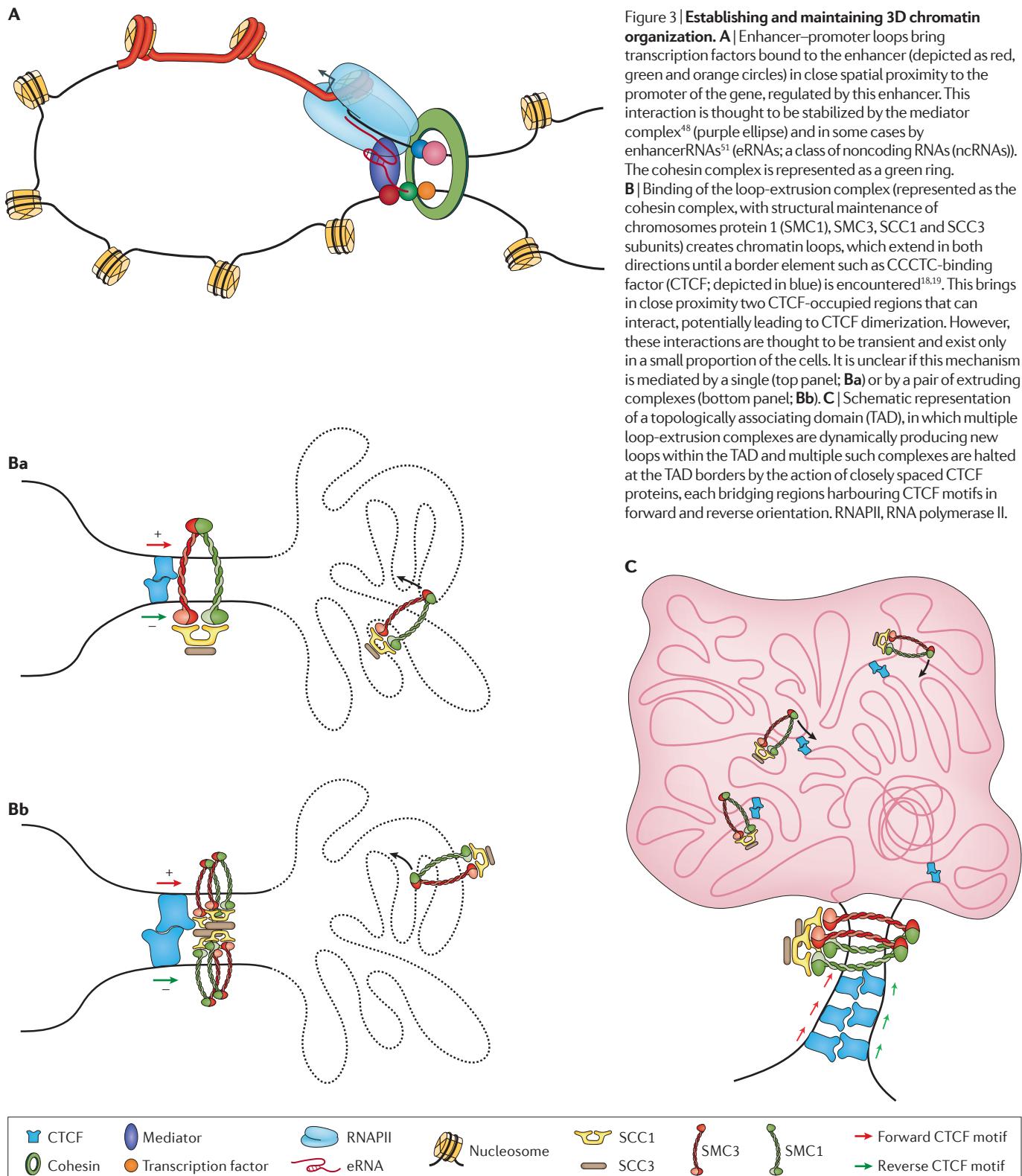
### Non-coding RNAs

One interesting observation is that both mediator and CTCF seem to be able to bind directly to RNA. In the case of mediator, MED12 (and to a lesser extent MED1) was found to bind specifically to non-coding RNAs (ncRNAs) called activating ncRNAs<sup>51</sup> (also known as enhancer RNAs (eRNAs)). Knockdown of these eRNAs led to a decrease in binding of mediator to genes regulated by the ncRNA as well as diminished loop formation between the ncRNA locus and its targets<sup>51,74</sup>. CTCF was also recently found to directly bind a large range of ncRNAs genome-wide<sup>75,76</sup>. CTCF contains an RNA-binding domain within its carboxyl-terminus, and multimerization of CTCF seems to depend on the presence of RNA, which has strong implications for chromatin topology<sup>76</sup>. Furthermore, YY1 — a ubiquitously expressed transcription factor shown to bind to CTCF<sup>77</sup> and enriched in chromatin loop anchors<sup>11</sup> — was also recently shown to bind RNA, which was suggested to reinforce transcription factor binding at regulatory elements<sup>78</sup>. However, it is unclear to what extent, if any, long ncRNAs (lncRNAs) contribute to the binding of CTCF or YY1. However, examples of ncRNAs regulating chromatin architecture are not limited to architectural proteins. The lncRNA

X chromosome inactivation  
Dosage compensation  
mechanism in mammals in  
which one of a pair of  
X chromosomes is silenced.

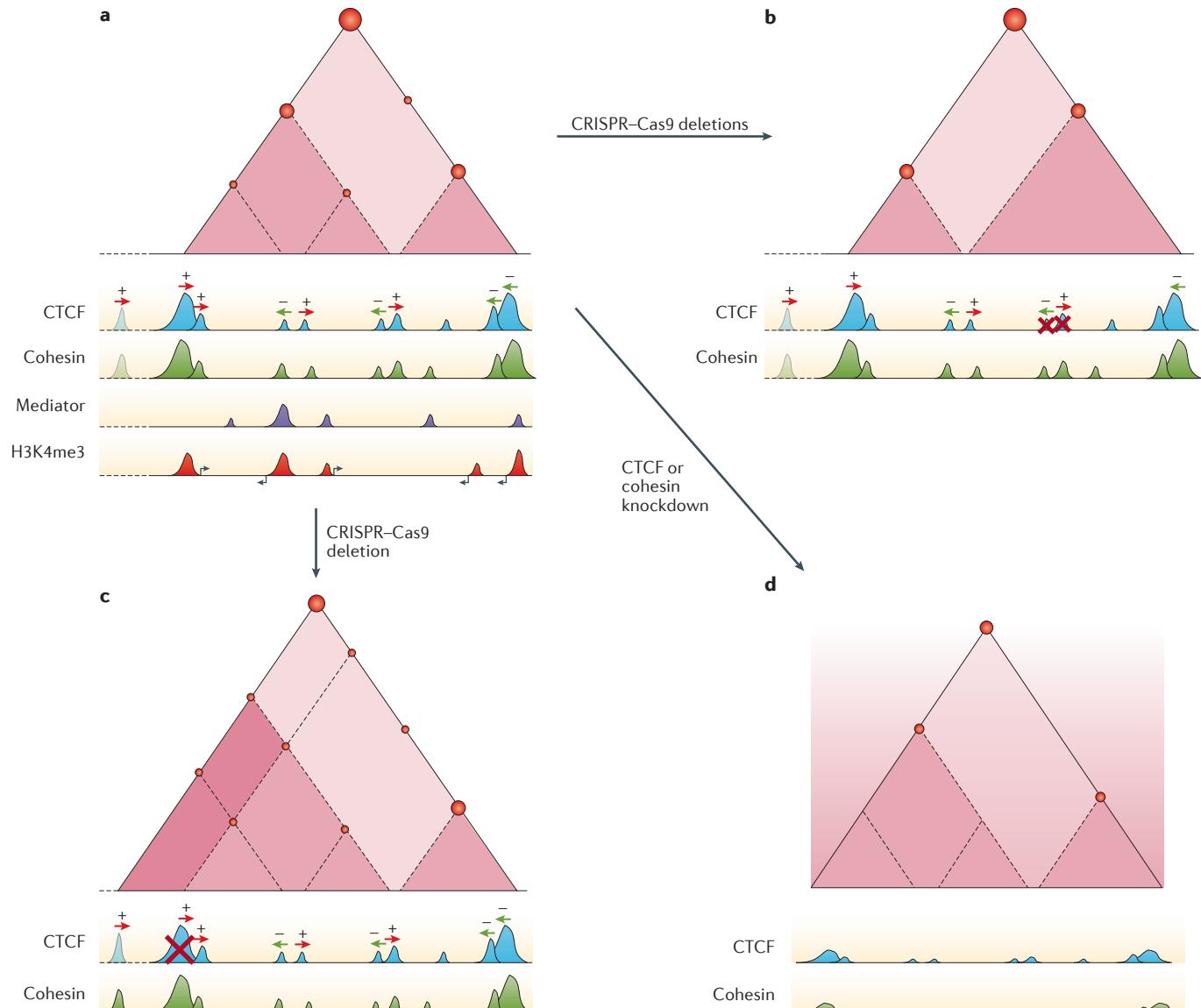
Firre (functional intergenic repeating RNA element) was shown to mediate the colocalization of several genomic regions, located on different chromosomes<sup>79</sup>. In the classic example of X chromosome inactivation, the ncRNA Xist is able to exploit 3D chromatin organization in order to

coat and mark one of the X chromosomes for inactivation<sup>80,81</sup>. Future work is required to dissect the precise molecular mechanisms at play and to determine whether establishing and maintaining 3D chromatin structure is a general role of nuclear lncRNAs.



**Combined effect of architectural components.** How different architectural proteins (and potentially ncRNAs) interact together to organize chromatin in 3D is still a matter of considerable debate. One prominent hypothesis recently put forth involves the combined action of loop-extrusion ‘motors’ (probably cohesins), which can dynamically bind and translocate chromatin to form a loop, until their progress on the chromatin fibre is halted by a ‘border element’ (proposed to be CTCF bound in a

specific orientation)<sup>18,19</sup> (FIG. 3b). This model is attractive because it can be used to explain the nesting of domains and loops on the basis of the assembly of possible states within a population and the consequences of CTCF motif deletion or inversion for loop and/or domain formation; it is also consistent with changes in 3D chromosome architecture observed in cohesin-depleted or CTCF-depleted cells. However, it is unclear whether the contact between loop anchors is dynamic or static



**Figure 4 | Importance of CTCF polarity on 3D chromatin organization.** **a** | Schematic representation of a typical contact domain, demarcated by a strong chromatin loop between the domain boundaries (red circle). Notice that several loops are also present within the topologically associating domain (TAD), leading to the formation of nested TAD-like structures (also known as subTADs<sup>36,37</sup>; demarcated by dotted lines). Architectural loops (demarcated by smaller circles) are usually formed between regions containing convergent CCCTC-binding factor (CTCF) motifs<sup>11,72</sup>. Active genes, demarcated by H3K4me3 enrichment and cohesin are also frequently found at domain boundaries<sup>9</sup>. **b** | Deletion of two of the CTCF-binding sites (indicated by red crosses) within the TAD leads to a change in intra-TAD

contacts and the emergence of novel chromatin loops<sup>19,72</sup>. **c** | Deletion of a CTCF loop located at the boundary of a TAD leads to an expansion of the domain to the closest upstream CTCF-binding site with a motif in the same orientation. **d** | Knockdown of CTCF leads to an increase in inter-TAD interactions and a decrease in intra-TAD contacts; however, TADs can still be recognized<sup>63</sup>. Intra-domain contacts are also disrupted upon cohesin depletion<sup>53,54</sup>. In all panels representative schematic Hi-C and chromatin immunoprecipitation sequencing (ChIP-seq) binding profiles for CTCF, cohesin, mediator and H3K4me3 are depicted to reflect TAD architecture. Different shades of red represent interaction strength between two regions. CRISPR, clustered regularly interspaced short palindromic repeats.

and what would be the consequence of biological differences in the properties of such hypothetical loop-extrusion enzymes, for example, during development or disease. Furthermore, it will be important to investigate whether chromatin loops can also be formed in other ways — perhaps by bulky multiprotein complexes such as RNAPII also acting as boundary elements.

### Implications of chromatin dynamics

Although the primary domain architecture of chromatin seems to be mainly preserved in different cell types and across species<sup>7,9,11</sup>, chromatin dynamics contribute to the specification of distinct gene expression programmes and biological functions. The mechanisms regulating dynamic chromatin changes are under intense investigation.

**Global chromatin reorganization.** Using diploid Hi-C maps, pronounced differences in chromatin architecture were observed between the active and inactive X chromosome in human<sup>11</sup> and mouse cells<sup>82</sup>. Whereas normal TAD structure was observed on the active X chromosome, two large domains, called ‘superdomains’ were identified in both species on the inactive X chromosome<sup>11,82</sup> (FIG. 5a). Importantly, whereas the genes located in these two superdomains differ between mouse and human, the border between them does not and is located near the macrosatellite large tandem repeat DXZ4, which encodes an ncRNA. On the active X chromosome of females and the X chromosome of males, DXZ4 is heterochromatic and does not bind CTCF, whereas on the inactive X chromosome in females it is euchromatic and binds CTCF<sup>83</sup>. Recently, the DXZ4 region was shown to be crucial for the formation of these two superdomains during X inactivation as well as the fine tuning of inactive X chromosome (Xi) chromatin function<sup>84,85</sup>. The role of chromatin organization in dosage compensation seems to be more general, as both the lncRNA Xist in mammals<sup>80,81</sup> and the male-specific lethal (MSL) complex in *D. melanogaster*<sup>86</sup> exploit the 3D organization of the X chromosome in order to spread, which enables Xist to mediate X inactivation in females and MSL to mediate transcriptional upregulation from the single X in males. In addition, a condensin-dependent architecture of the X chromosome, distinct from that of autosomes, was recently identified in *Caenorhabditis elegans*<sup>87</sup>.

Other dynamic processes have been shown to affect chromatin architecture at a global scale. During mitosis, chromosomes are strongly compacted and there is a widespread displacement of sequence-specific and basal transcription factors. The topological organization of the chromatin was shown to undergo a dramatic reorganization in M-phase and to be restored in early G<sub>1</sub> phase<sup>16</sup> (FIG. 5b), a finding that raises the question of how the pattern of 3D organization is re-established with each cell division. Furthermore, terminally differentiated post-mitotic cells often differ in their level of chromosomal rearrangements, which may be related to their different functions. For example, plasma cells have a smaller nucleus with a higher proportion of heterochromatin than dividing B cells<sup>88</sup>; rod photoreceptor cells in nocturnal

mammals have an unusual, inverted nuclear architecture, in which heterochromatin is localized in the centre of the nucleus and is absent from the periphery, an organization dependent on lamins A and C, (splice variants encoded by *LMNA*) and lamin B receptor (LBR)<sup>89</sup>; and mature neurons have elevated levels of polyploidy and aneuploidy<sup>90</sup>. All of these examples suggest that the requirement to undergo cell division in proliferating cells may limit the degree of freedom for changes in 3D architecture to occur; however, in postmitotic cells chromatin may be less constrained and may adopt a range of specialized structures to guide or to accompany cell function.

Supporting this hypothesis, two biological processes related to cell cycle exit have been shown to strongly affect chromatin 3D organization. During quiescence in yeast, intrachromosomal contacts increase, which is indicative of chromosome condensation, centromeres become more loosely associated and telomere interactions increase<sup>91</sup>. In senescence (characterized by irreversible cell cycle exit in response to exogenous and endogenous stress), heterochromatin relocates from the nuclear periphery to the interior, in some cases (when senescence is induced by an oncogene) forming nuclear structures known as senescence-associated heterochromatin foci (SAHF). This phenomenon is often accompanied with loss of lamin B1, which may function to anchor heterochromatin to the nuclear periphery<sup>92,93</sup>. Whereas the global domain structure remains mostly intact, local intra-TAD interactions seem to decrease in a sequence- and lamin-dependent manner and long-range contacts increase<sup>94</sup>.

**Smaller scale chromatin reorganization.** In contrast to these global changes, subtler effects are observed during biological processes such as differentiation and signalling. One recent study looked at changes in chromatin conformation during the transition from ES cells grown in 2i medium (which maintains ground-state pluripotency) to serum (in which ES cells become primed for differentiation)<sup>95</sup>. The authors discovered a gradual and reversible establishment of long-range interactions involving H3K27me3-marked bivalent promoters and Hox genes during the 2i-to-serum transition, which was dependent on the presence of Polycomb repressive complex 2 (PRC2)<sup>95</sup>. The role of Polycomb (and specifically of the Polycomb complex PRC1) in the formation of long-range contacts between gene promoters was further underscored by the disruption of Hox gene contacts in ES cells in which PRC1 component proteins RING1A and/or RING1B were deleted<sup>90</sup>. Thus, Polycomb complexes have an important role in organizing the 3D genome in early development (FIG. 5c), similarly to what was previously observed in *D. melanogaster*<sup>27</sup>.

**Chromatin reorganization during cell differentiation.** How does chromatin organization change during lineage specification? To address this question, a recent study examined 3D nuclear architecture in human ES cells and four different ES-derived lineages representing early developmental stages<sup>10</sup>. In agreement with previous results<sup>8,9</sup>, they found that the topological organization of the genome is mostly unchanged during

#### Boundary elements

DNA or epigenetic elements that are localized between two topological domains and that prevent or minimize inter-domain interactions.

#### Dosage compensation

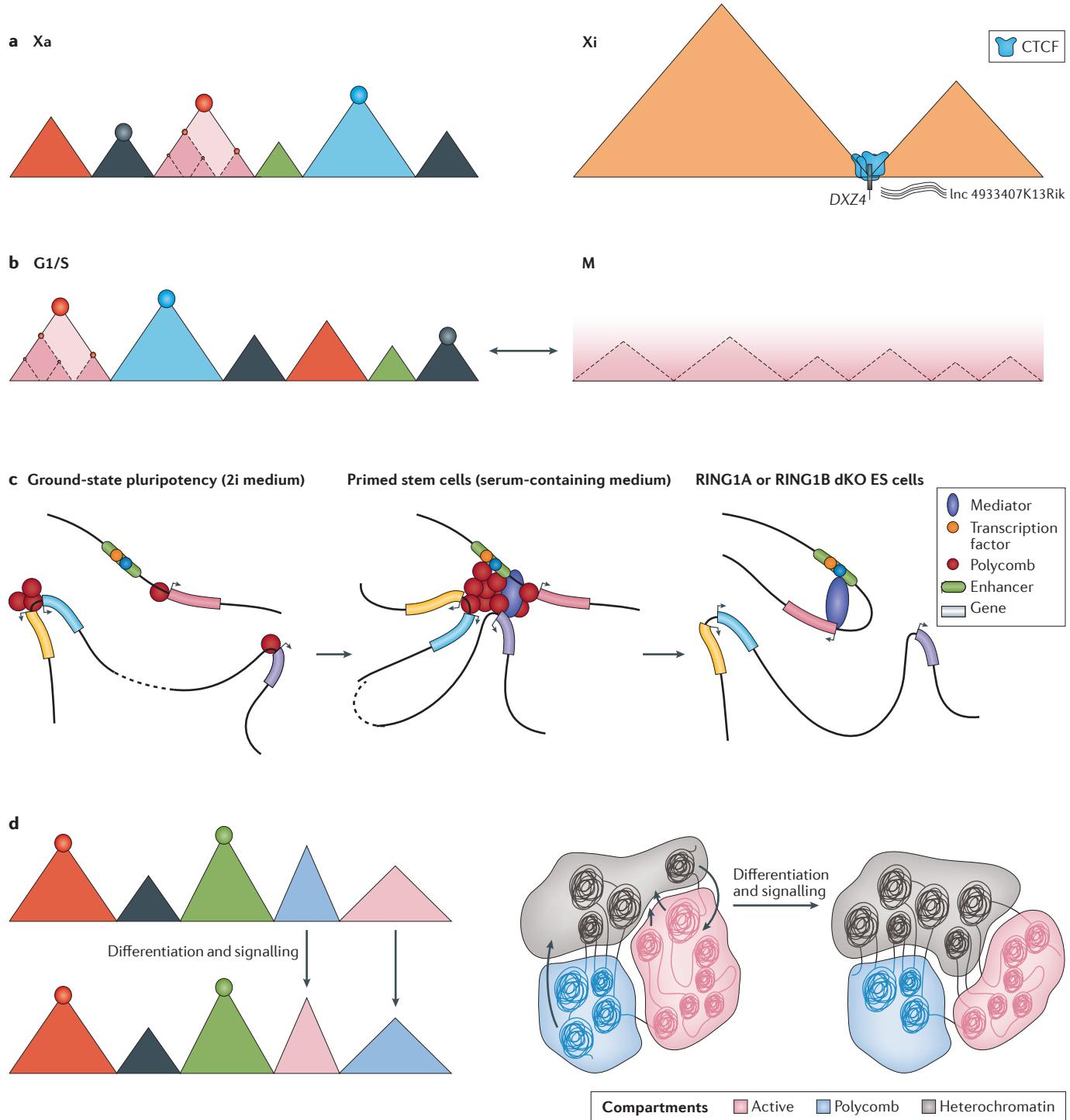
The process of equalizing expression output from genes located on the sex-specific chromosomes.

#### Polyploidy

An increase in the number of chromosomes in a cell by whole-number multiples of the entire set.

#### Aneuploidy

Aberrations in the number of chromosomes, usually accompanied by structural rearrangements.



**Figure 5 | Static and dynamic components of chromatin organization.** **a** | Three-dimensional (3D) organization of the X chromosome in mouse and human. Notice that although the active X chromosome (Xa) has normal topological organization, there are only two ‘superdomains’ present on the inactive X chromosome (Xi)<sup>11,84</sup>. Circles represent chromatin loops. DXZ4 refers to a repeat region on the X chromosome that binds CCCTC-binding factor (CTCF) and produces a long non-coding RNA (4933407K13Rik) only on the inactive X chromosome. **b** | Topologically associating domain (TAD)-like organization of the genome is proposed to be lost during mitosis (denoted by M; right panel) during chromosome condensation and re-established in early G1 to S (left panel) phase<sup>10</sup>. Circles represent chromatin loops. **c** | Polycomb-mediated long-range contacts in embryonic stem (ES) cells are established during the transition between ground-state (2i medium; left) and primed (serum-containing medium; middle)<sup>95</sup> and are lost in cells lacking the Polycomb repressor complex 1 (PRC1) subunits RING1A and RING1B (right)<sup>30</sup>. **d** | During differentiation and upon external stimuli TADs can acquire different chromatin marks (left panel) and shift between different compartments (right panel)<sup>10</sup>. Circles represent chromatin loops. dKO, double knockout.

lineage specification, but intra-TAD interactions in some domains were strongly altered and the direction of these changes correlated positively with an open chromatin state<sup>10</sup>. This TAD-wide change in interactions often correlated with a relocation of the TAD from one compartment to another and with changes in the transcription status of the genes belonging to the TAD<sup>10</sup> (FIG. 5d). However, only changes in early developmental lineages were examined, so it will be interesting to analyse how chromatin contacts change in a gradual, multi-stage, cell cycle-matched differentiation system.

Another study examined the nuclear architecture in B cell differentiation. Several regions were shown to switch compartment identity and, in the case of the early B cell factor 1 (*EBF1*) locus, to relocate from the nuclear periphery to the nuclear interior<sup>96</sup>. Furthermore, loops anchored by E1A-binding protein (also known as p300) or the lineage-specific transcription factors E2A and PU.1 were found to be developmentally regulated<sup>96</sup>, suggesting that transcription factors are capable of rearranging chromatin architecture.

To study how chromatin architecture responds to transient stimuli, such as hormone signalling, the effect of treatment with progestin or estradiol on 3D nuclear structure in breast cancer cells was examined<sup>97</sup>. Despite large changes in the transcriptional output of these cells, only small changes were observed in the topological organization of chromatin, with only a few dynamic boundary regions. However, for a substantial number of domains, the entire TAD responded to the hormone treatment as a unit, by changing the epigenetic signature and switching between the A and B compartment, which suggests that transcription status is coordinated within a TAD<sup>97</sup>. However, in these and other studies it is unclear whether changes in transcription and/or chromatin marks are a cause or a consequence of changes in genome architecture.

**Interplay between transcription and chromatin looping.** Originally, during the study of the formation of the  $\beta$ -globin active chromatin hub, long-range interactions were proposed to form in cells in which the target gene is active<sup>25</sup>, presumably because of tissue-specific factors. However, recent evidence in mice suggests that this might not be the case and that transcriptional output can be, at least temporally, uncoupled from chromatin connectivity. In posterior limb buds the expression of the sonic hedgehog protein (*Shh*) is regulated by a distal enhancer called ZRS (zone of polarizing activity regulatory sequence), which forms a chromatin loop and contacts the *Shh* locus<sup>98</sup>. This loop seems to be preset and it is detectable even where *Shh* is not transcribed, such as in anterior limb buds<sup>98</sup>. Analogous results were obtained when examining the regulatory sequences within the HoxD cluster. These elements were found to contact each other and the target genes to form a hub, and some of the interactions were present even in cells in which the target genes were not transcribed<sup>99,100</sup>. Consistent with these results, a large number of enhancer–promoter interactions seem to be stable, associated with paused RNAPII and preset before

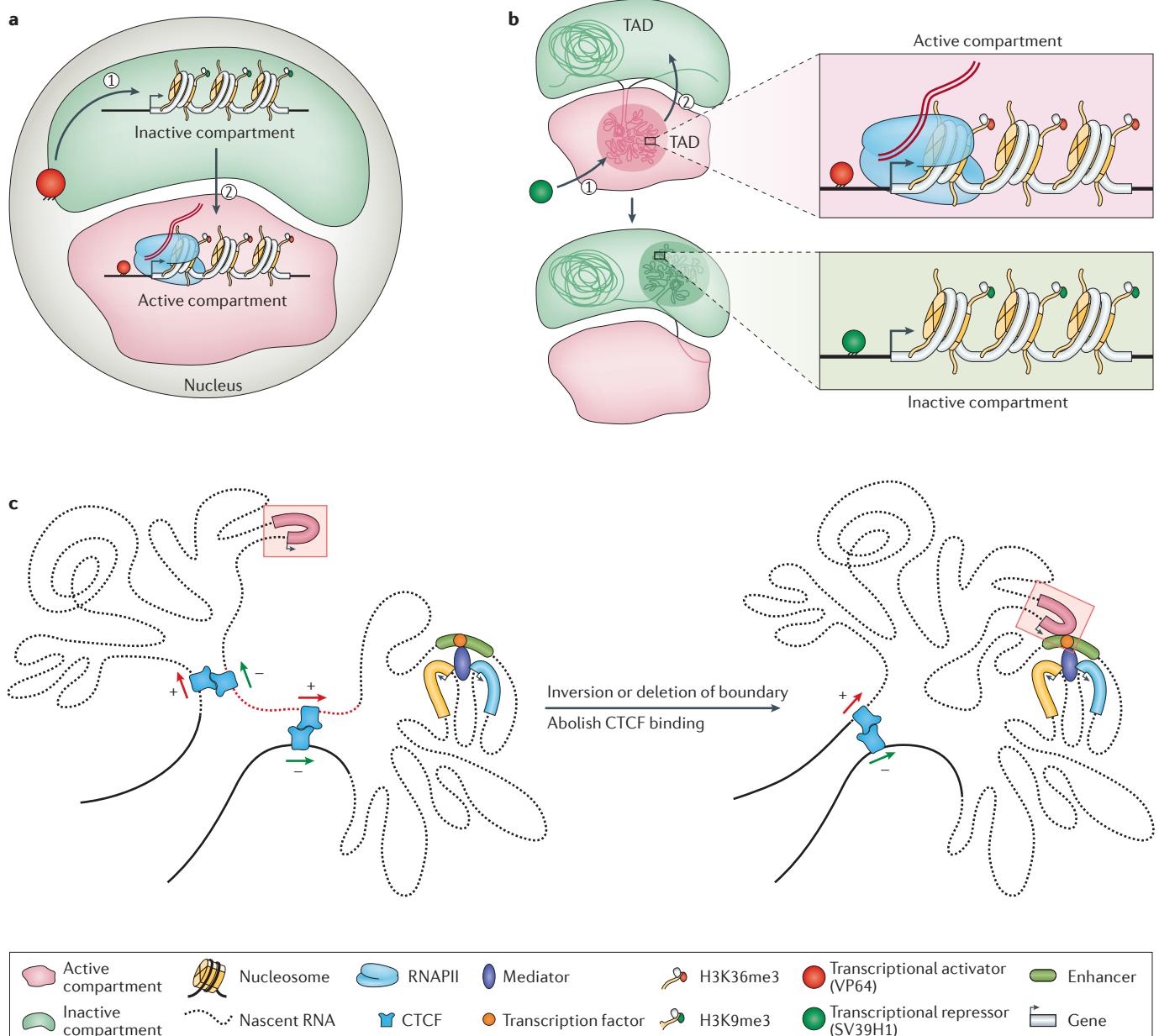
gene activation during *D. melanogaster* development<sup>101</sup>. These results suggest that the release of RNAPII from pausing is crucial for tissue-specific gene activation, not the formation of an enhancer–promoter loop. It will be important to confirm these results in mammalian systems and to extend them genome-wide, for example, using high-resolution Hi-C.

### 3D organization and gene expression

Gene positioning within the 3D nuclear organization depends on the chromatin status as well as the transcriptional output of the locus. Gene-dense chromosomes and chromosomal regions are located predominantly within the euchromatic interior of the nucleus, whereas gene-poor, heterochromatic and late-replicating domains are found close to the nuclear envelope. This radial positioning has been shown to be dependent on either LBR or lamins A and C, as the absence of these components led to an accumulation of heterochromatin at the nuclear centre<sup>89</sup>. Indeed, DNA adenine methyltransferase identification (DamID) analysis, has shown that ~40% of the genome is engaged in the formation of so-called lamina-associated domains (LADs) in human fibroblasts<sup>102</sup>. These LADs are generally gene-poor and associated with low levels of gene expression. However, gene positioning within the nuclear environment is not always fixed and the actual association of LADs with the nuclear lamina is not constant even within the same cell type<sup>103</sup>. Only 30% of the LADs do contact the lamina in any given cell and they seem to randomly attach or detach at every cell cycle<sup>103</sup>. Furthermore, DamID confirmed previous observations by FISH that during the differentiation of mouse ES cells, LADs can be at least partially dynamic and cell-type specific<sup>104</sup>. A loss of a lamina interaction in an intermediate stage of differentiation can poised the locus for activation during subsequent differentiation stages. However, these observations did not determine whether differential gene expression is causative or a consequence of relocalization relative to the nuclear periphery.

**Disentangling cause and consequence.** The work of several laboratories has recently shown that chromatin decondensation alone (without activating transcription) is sufficient to cause relocation of a locus from the nuclear periphery towards to the centre<sup>4</sup> (FIG. 6a). Furthermore, the knockdown of specific transcription factors and chromatin remodelers such as structural maintenance of chromosomes 3 (SMC3) and SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member (SMARCD2) is sufficient to cause a relocation of some highly active genes towards the nuclear periphery without affecting their expression levels<sup>105</sup> and this process is dependent on progression through DNA replication but not mitosis. In support of the uncoupling between the transcriptional output of a gene and its location within the nucleus, *C. elegans* chromodomain protein (CEC-4) was identified as being necessary for the anchoring of heterochromatin to the nuclear lamina without affecting its transcription status<sup>5</sup>.

DNA adenine methyltransferase identification (DamID). Technique to identify the binding sites of DNA- and chromatin-binding proteins in eukaryotes by fusing them to the bacterial methyltransferase enzyme Dam.



**Figure 6 | 3D genome organization and gene expression.** **a** | Artificial recruitment of a transcriptional activator (such as VP64, depicted by the red circle) or chromatin decondensation alone is sufficient to reposition a locus located normally in the nuclear periphery towards the nuclear centre<sup>4</sup>. **b** | Artificial tethering of a transcriptional repressor (such as SUV39H1, depicted by the green circle) to an active locus (1) shifts the whole sub-topologically associating domain (subTAD) containing this locus to the nuclear periphery (2)<sup>37</sup>. **c** | Absence of boundary elements caused by genetic (deletion or inversion) or epigenetic mechanisms (DNA methylation) can have consequences for gene expression, for example, by bringing an active enhancer located in one TAD (green) in close proximity to a normally inactive gene (pink), leading to aberrant transcription of the gene. 3D, three-dimensional; CTCF, CCCTC-binding factor; RNAPII, RNA polymerase II.

These and previous examples show that nuclear architecture is correlated with and underlies gene expression, but the phenotypical consequences of altering 3D genome organization are not well understood. In a landmark study, forcing a loop between the  $\beta$ -globin promoter and the locus control region (LCR) in the absence of the transcription factor GATA1 (also known as erythroid transcription factor), which is normally required for  $\beta$ -globin expression, was sufficient

to recruit RNAPII and to substantially upregulate the expression of the  $\beta$ -major globin gene<sup>106</sup>. This study showed for the first time that chromatin looping alone is sufficient to activate gene expression. Furthermore, an engineered chromatin loop between the LCR and the  $\gamma$ -globin promoter in adult human erythrocytes led to the upregulation of fetal-stage  $\gamma$ -globin transcription to ~85% of total globin levels at the expense of adult  $\beta$ -globin transcription<sup>107</sup>, showing that chromatin

interactions can have an instructive role in gene expression and can determine the outcome of developmental choices. However, as discussed in the previous section, most enhancer–promoter contacts seem to be preset in *D. melanogaster* before gene activation<sup>101</sup>, so it is unclear whether looping alone can account for RNAPII activation and transcription globally. One possibility is that looping, established by specific machineries in order to set the permissive condition for gene activation, can be followed by actual activation either immediately or at later time points, depending on regulatory cues. Artificial recruitment of transcription factors (for example, NANOG) or chromatin modifiers (for example, EZH2 or SUV39H1) to different genomic loci showed that entire TADs can be repositioned to a different subcompartment or, in the case of SUV39H1 recruitment, can switch from the active A to the inactive B compartment<sup>37</sup> (FIG. 6b). Such repositioning seems to be uncoupled from transcriptional changes, which is consistent with the findings of previous studies<sup>4,5,105</sup>. Importantly, in the case of SUV39H1 recruitment, repositioning of the locus depends on the presence of the chromodomain of SUV39H1 and not on the enzymatic activity of the protein or the H3K9me3 mark deposited by it<sup>37</sup>.

**Chromatin architecture in development and disease.** A recent study examined how structural variation in the human genome, such as limb phenotypes associated with large-scale inversion, deletions and duplications within the *WNT6*–Indian hedgehog (*IHH*)–ephrin type A receptor 4 (*EPHA4*)–paired box 3 (*PAX3*) locus, can affect gene expression and can cause pathogenic phenotypes<sup>108</sup>. All of these structural changes disrupted a TAD boundary within the above-mentioned locus and led to ectopic interactions between a cluster of limb enhancers normally confined to the *EPHA4* TAD and gene promoters located outside of it. This was shown to depend on the CTCF-associated boundary elements<sup>108</sup>. Could there be other potential mechanisms, perhaps epigenetic, that would permit ectopic interactions between regions in two adjacent TADs? In gliomas associated with mutations in isocitrate dehydrogenase (*IDH*) genes in which DNA methylation levels are globally increased, CTCF sites located at a TAD boundary region close to the glioma oncogene platelet derived growth factor receptor- $\alpha$  (*PDGFRA*) become methylated, which the authors propose leads to decreased binding of CTCF to the boundary and to ectopic activation of the *PDGFRA* gene by an enhancer located in the adjacent TAD<sup>109</sup>. This is dependent on the CTCF-binding site within the boundary, as CRISPR (clustered regularly interspaced short palindromic repeats)-mediated deletion of the CTCF site or treatment with a DNA-demethylating agent (5'azacytidine) had similar effects, showing that the ectopic interaction was reversible<sup>109</sup>. However, as discussed previously, the causal relationship between DNA methylation and CTCF binding is still a matter of a debate and the observed effect of 5'azacytidine on the CTCF site at the *PDGFRA* locus was relatively small

(1.7-fold increase)<sup>109</sup>. In an analogous study, deletions associated with anchors of strong chromatin loops or domain boundaries were shown to be frequent in cancer, often leading to upregulation of a proto-oncogene enclosed within the loop or domain<sup>110</sup>. These studies suggest that ectopic inter-TAD contacts can occur when CTCF binding at boundaries is abrogated or diminished, and in some cases novel loops can lead to misexpression of important genes and severe phenotypical consequences (FIG. 6c). It will be interesting to examine how changes in CTCF binding during development, perhaps in relation to DNA methylation levels, would globally affect the rearrangement of dynamic enhancer–promoter interactions.

To further understand the role of genome architecture in development and disease, it is important to examine its contribution to the regulation of chromatin states and transcription within a population. Two papers addressed how genetic variation is associated with changes in enhancer marks, chromatin accessibility and transcription (quantitative trait loci (QTL))<sup>111,112</sup>. Although local single nucleotide polymorphisms (SNPs) in regulatory regions affected chromatin states and gene expression locally as expected, they were also more coordinated with changes in the chromatin status of physically interacting distal QTLs (>50 kb away) compared with non-interacting loci. Furthermore, distal QTLs seem to be enriched within TADs, changes in chromatin state occur concordantly between them and local-distal QTL pairs predominantly involve pairs of enhancers<sup>111</sup>. This is consistent with the idea of chromatin hubs<sup>25,113,114</sup>, in which several regulatory regions are physically connected with their target genes and can elicit a coordinated response. In this model, a change in the chromatin state of one such element, for example because of a genetic variation within a transcription factor-binding site, can modulate the epigenetic marks of the proximal (in 3D space) regions.

### Chromatin organization and evolution

Given the contribution of nuclear architecture to gene expression, it is important to consider how 3D organization can affect genome evolution. Topological organization of the genome into TADs has been observed in *D. melanogaster* and mammals, but how common are such structures in other species?

Although not initially observed in budding yeast<sup>115</sup>, recent nucleosome-resolution chromatin-interaction maps uncovered domain-like structures (called chromosomal interacting domains (CIDs)), which are much shorter than the megabase scale TADs in mammals and generally encompass one to five genes<sup>23</sup>. Similar topological organization and enrichment of active genes at boundaries was also observed in the genome of the bacteria *Caulobacter crescentus*<sup>116</sup>. Self-interacting domains (SIDs) with an average size of ~50–100 kb, called ‘globules’, were also observed in the fission yeast *Schizosaccharomyces pombe* and their formation was found to be dependent on cohesin<sup>117</sup>, whereas SIDs with an average size of ~1 Mb were observed only on the X chromosome in *C. elegans*<sup>87</sup>. In plants, the existence of

Quantitative trait loci (QTL). Regions in the genome that correlate with phenotypic variation.

TAD-like structures is still a matter of debate. In one study, very few small interactive regions were found in *Arabidopsis thaliana* and those were primarily enriched in repressive marks such as H3K27me3 and H3K9me2 (REFS 118,119), whereas in another study, large domain-like structures called 'structural domains' were observed<sup>119</sup>. Although it is clear that chromosome domains exist in a large range of species, further studies providing chromatin contact maps with higher spatial resolution will be required to establish whether they are a fundamental and obligatory feature of eukaryotic genomes.

The conservation between 3D organization in different species extends beyond domains. In particular, syntetic regions between mouse and human seem to have a more conserved 3D organization, indicating that similarity is not limited to the linear sequence<sup>9</sup>, an observation that was later validated and extended to four different mammalian species<sup>70</sup>. This was shown to be dependent on the conservation of strong CTCF sites, which colocalize with cohesin and determine the conserved TAD boundaries<sup>70</sup>. Furthermore, distant human loci that were adjacent in the mouse genome retained chromatin contacts more often than expected after they became separated on the linear genome through evolutionary rearrangements<sup>120</sup>, and long-range contacts between Hox loci, which are mediated by Polycomb, were conserved in fly species that diverged 40 million years ago<sup>27</sup>. Chromatin architecture also influences genomic rearrangements during evolution. For example, both evolutionary and disease-originating break points are distributed non-randomly in the genome and tend to occur more frequently in regions characterized by high gene density, high GC content and mostly open chromatin<sup>121–123</sup>. Ancestral genome reconstruction and statistical modelling showed that observed rearrangements can be accurately reproduced by taking into account the 3D nuclear organization<sup>124</sup>. The authors suggest that chromosomal rearrangements are more likely to occur between double-stranded DNA breaks in active chromatin domains that are in close spatial proximity to each other<sup>124</sup>. It will be important to carry out further studies to investigate this hypothesis and to extend it to specific evolutionary events.

### Perspectives

Only a decade after the advent of molecular biology methods to study chromatin contacts at the genome-wide level, it has become clear that 3D genome architecture is intimately linked to regulating gene expression during development, in physiological processes and in disease. The discovery of epigenomic chromosomal domains and of TADs has added a new dimension to our understanding of genome function and most recent analyses in the field have been directed towards understanding TAD formation and function. A future challenge will be to extend the analysis to the larger and more elusive chromosome compartments: studying in which species they exist, their evolutionary role, how they are formed and their role in gene regulation and in other DNA-dependent processes, such as DNA

replication, recombination and repair. Improving our understanding and our ability to predict the outcome of architectural genome changes and how these could be modulated for therapy will require further technological developments, which are underway. 3C-type methods will further improve, both in the sequencing depth and by refinement of the current approaches. Single cell 3C-based approaches may provide information on cell-to-cell architectural variability, but cannot describe chromatin dynamics. Microscopy is greatly improving and, just as the evolution of 3C into Hi-C has provided a new dimension in the molecular understanding of the 3D genome, development of conventional microscopy into a 'Hi-M' methodology that may combine high-throughput ultra-fast image acquisition with super-resolution microscopy will bring us to a new dimension of high-resolution image-omics data. Further improvements of current live imaging may allow tracking of the dynamics of chromatin domains and interactions in live cells in order to investigate conformational changes upon various stimuli and in relation to gene expression. These complex multi-dimensional data call for extensive quantitative analyses, and computational biology is developing in this direction. Mathematical modelling can complement biological investigation and rationalize as well as predict important aspects of chromatin behaviour.

As a note of caution, although microscopy-based methods are usually in good agreement with 3C approaches<sup>8,44,125</sup> in some cases (for example, Hox genes in Polycomb mutants) the conclusions reached using different methods are not always in agreement<sup>126</sup>. Such inconsistencies suggest that we need to invest energy in assessing the limitations and possible caveats of experimental approaches, in order to correct biases and to improve convergence between them. In addition, most of the studies of genome architecture reported so far were generated using cell lines or heterogeneous tissues and may not reflect chromatin architecture *in vivo*. For example, cells cultured *in vitro* have been shown to have a higher proportion of heterochromatic regions compared with primary cells<sup>127</sup>, which will probably reflect the chromatin conformation. Further efforts to scale down cell numbers needed for 3C-based methods in order to generate chromatin interaction maps from pure, fluorescence-activated cell sorting (FACS)-purified populations *in vivo* will be required on the molecular side and improvement of high-resolution microscopy methods allowing the study of cells in tissues will be essential on the imaging front.

Nevertheless, even with imperfect methodology, we have observed an unprecedented boom in our understanding of chromosome folding and its relation to function. However, this represents only the tip of the iceberg of chromatin biology and the next few years will probably lead to unanticipated insights about the molecular mechanisms behind the establishment and the maintenance of the 3D genome, the relationship between genome organization and transcription, and the importance of chromatin architecture for normal development, disease and evolution.

In summary, the simultaneous development of technological and scientific approaches is leading us to an integrated understanding of the function of the genome and its associated components in development, physiology and disease. The combination of

these tools with functional studies, particularly those made possible by the advent of genome-engineering technologies such as CRISPR-Cas9 (REF. 128), promises to lead to major advances for this novel field in the near future.

1. Bickmore, W. A. & van Steensel, B. Genome architecture: domain organization of interphase chromosomes. *Cell* **152**, 1270–1284 (2013).
2. Sexton, T. & Cavalli, G. The role of chromosome domains in shaping the functional genome. *Cell* **160**, 1049–1059 (2015).
3. Pombo, A. & Dillon, N. Three-dimensional genome architecture: players and mechanisms. *Nat. Rev. Mol. Cell Biol.* **16**, 245–257 (2015).
4. Therizols, P. *et al.* Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science* **346**, 1238–1242 (2014). **In this paper, transcriptional activation or chromatin decondensation alone is sufficient to cause a translocation of the underlying locus from the nuclear periphery towards the nuclear core.**
5. Gonzalez-Sandoval, A. *et al.* Perinuclear anchoring of H3K9-Methylated chromatin stabilizes induced cell fate in *C. elegans* embryos. *Cell* **163**, 1333–1347 (2015).
6. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
7. Sexton, T. T. *et al.* Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012). **This paper describes the discovery of TADs in *Drosophila melanogaster* and shows that TADs overlap extensively with distinct patterns of epigenetic marks.**
8. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012). **In this study, the authors describe the discovery of TADs in the X chromosome using 5C and show that the boundaries of those TADs are defined by cis-acting genetic elements.**
9. Dixon, J. R. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012). **In this paper, the global organization of the mammalian genomes into TADs is reported and TAD boundaries are shown to be relatively constant between cell types and enriched in CTCF.**
10. Dixon, J. R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).
11. Rao, S. S. P. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014). **The authors use Hi-C to characterize the chromatin organization in nine different human and mouse cell lines with very high resolution. They show that chromatin loops are often established between two CTCF sites with convergent motif orientation.**
12. Schuettengruber, B. *et al.* Cooperativity, specificity, and evolutionary stability of Polycomb targeting in *Drosophila*. *Cell Rep.* **9**, 219–233 (2014).
13. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
14. Giorgetti, L. *et al.* Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* **157**, 950–963 (2014).
15. Nagano, T. *et al.* Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2014). **In this paper, the authors use single cell Hi-C to examine the heterogeneity of 3D genome organization within a population of cells.**
16. Naumova, N. *et al.* Organization of the mitotic chromosome. *Science* **342**, 948–953 (2013).
17. Marsden, M. P. & Laemmli, U. K. Metaphase chromosome structure: evidence for a radial loop model. *Cell* **17**, 849–858 (1979).
18. Fudenberg, G. *et al.* Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
19. Sanborn, A. L. *et al.* Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. USA* **112**, E6456–E6465 (2015).
20. Schalch, T., Duda, S., Sargent, D. F. & Richmond, T. J. X-Ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436**, 138–141 (2005).
21. Tremethick, D. J. Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell* **128**, 651–654 (2007).
22. Ricci, M. A., Manzo, C., García-Parajo, M. F., Lakadamyali, M. & Cosma, M. P. Chromatin fibers are formed by heterogeneous groups of nucleosomes *in vivo*. *Cell* **160**, 1145–1158 (2015).
23. Hsieh, T.-H. S. *et al.* Mapping nucleosome resolution chromosome folding in yeast by micro-C. *Cell* **162**, 108–119 (2015).
24. Fussner, E. *et al.* Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres. *Cell* **13**, 992–926 (2012).
25. Palstra, R. J. *et al.* The β-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.* **35**, 190–194 (2003).
26. Schoenfelder, S. *et al.* Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet.* **42**, 53–61 (2010).
27. Bantignies, F. *et al.* Polycomb-dependent regulatory contacts between distant Hox loci in *Drosophila*. *Cell* **144**, 214–226 (2011).
28. Denholtz, M. *et al.* Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and Polycomb proteins in genome organization. *Cell Stem Cell* **13**, 602–616 (2013).
29. Vieux-Rochas, M., Fabre, P. J., Leleu, M., Duboule, D. & Noordermeer, D. Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain. *Proc. Natl. Acad. Sci. USA* **112**, 4672–4677 (2015).
30. Schoenfelder, S. *et al.* Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. *Nat. Genet.* **47**, 1179–1186 (2015).
31. O'Sullivan, J. M. *et al.* Gene loops juxtapose promoters and terminators in yeast. *Nat. Genet.* **36**, 1014–1018 (2004).
32. Tan-Wong, S. M. *et al.* Gene loops enhance transcriptional directionality. *Science* **338**, 671–675 (2012).
33. Shen, Y. *et al.* A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116–120 (2012).
34. Eagen, K. P., Hartl, T. A. & Kornberg, R. D. Stable chromosome condensation revealed by chromosome conformation capture. *Cell* **163**, 934–946 (2015).
35. Ulianov, S. V. *et al.* Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. *Genome Res.* **26**, 70–84 (2016).
36. Phillips-Cremins, J. E. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281–1295 (2013).
37. Wijchers, P. J. *et al.* Cause and consequence of tethering a subTAD to different nuclear compartments. *Mol. Cell* **61**, 461–473 (2016).
38. Wang, S. *et al.* Spatial organization of chromatin domains and compartments in single chromosomes. *Science* **353**, 598–602 (2016).
39. Boettiger, A. N. *et al.* Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).
40. Lichter, P., Cremer, T., Borden, J., Manuelidis, L. & Ward, D. C. Delineation of individual human chromosomes in metaphase and interphase cells by *in situ* suppression hybridization using recombinant DNA libraries. *Hum. Genet.* **80**, 224–234 (1988).
41. Pinkel, D. *et al.* Fluorescence *in situ* hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. *Proc. Natl. Acad. Sci. USA* **85**, 9138–9142 (1988).
42. Van Bortle, K. *et al.* Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol.* **15**, R82 (2014).
43. Schwartz, Y. B. *et al.* Nature and function of insulator protein binding sites in the *Drosophila* genome. *Genome Res.* **22**, 2188–2198 (2012).
44. Giorgetti, L., Servant, N. & Heard, E. Changes in the organization of the genome during the mammalian cell cycle. *Genome Biol.* **14**, 142 (2013).
45. Heath, H. *et al.* CTCF regulates cell cycle progression of αβ T cells in the thymus. *EMBO J.* **27**, 2839–2850 (2008).
46. Allen, B. L. & Taatjes, D. J. The mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 155–166 (2015).
47. Nasmyth, K. & Haering, C. H. Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* **43**, 525–558 (2009).
48. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
49. Splinter, E. *et al.* CTCF mediates long-range chromatin looping and local histone modification in the β-globin locus. *Genes Dev.* **20**, 2349–2354 (2006).
50. Hadjur, S. *et al.* Cohesins form chromosomal *cis*-interactions at the developmentally regulated IFNG locus. *Nature* **460**, 410–413 (2009).
51. Lai, F. F. *et al.* Activating RNAs associate with mediator to enhance chromatin architecture and transcription. *Nature* **494**, 497–501 (2013).
52. Rubio, E. D. *et al.* CTCF physically links cohesin to chromatin. *Proc. Natl. Acad. Sci. USA* **105**, 8309–8314 (2008).
53. Seitan, V. C. *et al.* Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res.* **23**, 2066–2077 (2013).
54. Sofueva, S. *et al.* Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J.* **32**, 3119–3129 (2013).
55. Ong, C.-T. & Corces, V. G. CTCF: an architectural protein bridging genome topology and function. *Nat. Rev. Genet.* **15**, 234–246 (2014).
56. Kurukuti, S. *et al.* CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc. Natl. Acad. Sci. USA* **103**, 10684–10689 (2006).
57. Xie, X. *et al.* Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proc. Natl. Acad. Sci. USA* **104**, 7145–7150 (2007).
58. Heger, P., Marin, B., Bartkuhn, M., Schierenberg, E. & Wiehe, T. The chromatin insulator CTCF and the emergence of metazoan diversity. *Proc. Natl. Acad. Sci. USA* **109**, 17507–17512 (2012).
59. Soshnikova, N., Montavon, T., Leleu, M., Galjart, N. & Duboule, D. Functional analysis of CTCF during mammalian limb development. *Dev. Cell* **19**, 819–830 (2010).
60. Wan, L. B. *et al.* Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development. *Development* **135**, 2729–2738 (2008).
61. Narendra, V. *et al.* CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* **347**, 1017–1021 (2015).
62. Handoko, L. *et al.* CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat. Genet.* **43**, 630–638 (2011).
63. Zuin, J. *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. USA* **111**, 996–1001 (2014).
64. Chen, H., Tian, Y., Shu, W., Bo, X. & Wang, S. Comprehensive identification and annotation of cell type-specific and ubiquitous CTCF-binding sites in the human genome. *PLoS ONE* **7**, e41374 (2012).
65. Bell, A. C. & Felsenfeld, G. Methylation of a CTCF-

- dependent boundary controls imprinted expression of the *ICF2* gene. *Nature* **405**, 482–485 (2000).
66. Hark, A. T. *et al.* CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/ICF2* locus. *Nature* **405**, 486–489 (2000).
  67. Wang, H. *et al.* Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res.* **22**, 1680–1688 (2012).
  68. Stadler, M. B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495 (2011).
  69. Feldmann, A. *et al.* Transcription factor occupancy can mediate active turnover of DNA methylation at regulatory regions. *PLoS Genet.* **9**, e1003994 (2013).
  70. Vietri Rudan, M. *et al.* Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.* **10**, 1297–1309 (2015).
  71. Tang, Z. *et al.* CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* **163**, 1611–1627 (2015).
  72. de Wit, E. *et al.* CTCF binding polarity determines chromatin looping. *Mol. Cell* **60**, 676–684 (2015).
  73. Guo, Y. *et al.* CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* **162**, 900–910 (2015).
  74. Trimarchi, T. *et al.* Genome-wide mapping and characterization of Notch-regulated long noncoding RNAs in acute leukemia. *Cell* **158**, 593–606 (2014).
  75. Kung, J. T. *et al.* Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. *Mol. Cell* **57**, 361–375 (2015).
  76. Saldana-Meyer, R. *et al.* CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev.* **28**, 723–734 (2014).
  77. Donohoe, M. E., Zhang, L.-F., Xu, N., Shi, Y. & Lee, J. T. Identification of a CTCF cofactor, YY1, for the X chromosome binary switch. *Mol. Cell* **25**, 43–56 (2007).
  78. Sigova, A. A. *et al.* Transcription factor trapping by RNA in gene regulatory elements. *Science* **350**, 978–981 (2015).
  79. Hacisuleyman, E. *et al.* Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* **21**, 198–206 (2014).
  80. Engreitz, J. M. *et al.* The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**, 1237973 (2013).
  81. Simon, M. D. *et al.* High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* **504**, 465–469 (2013).
  82. Deng, X. *et al.* Bipartite structure of the inactive mouse X chromosome. *Genome Biol.* **16**, 152 (2015).
  83. Horakova, A. H., Moseley, S. C., McLaughlin, C. R., Tremblay, D. C. & Chadwick, B. P. The macrosatellite DXZ4 mediates CTCF-dependent long-range intrachromosomal interactions on the human inactive X chromosome. *Hum. Mol. Genet.* **21**, 4367–4377 (2012).
  84. Giorgetti, L. *et al.* Structural organization of the inactive X chromosome in the mouse. *Nature* **535**, 575–579 (2016).
  85. Darow, E. M. *et al.* Deletion of DXZ4 on the human inactive X chromosome alters higher-order genome architecture. *Proc. Natl. Acad. Sci. USA* **113**, E4504–E4512 (2016).
  86. Ramírez, F. *et al.* High-affinity sites form an interaction network to facilitate spreading of the MSL complex across the X chromosome in *Drosophila*. *Mol. Cell* **60**, 146–162 (2015).
  87. Crane, E. *et al.* Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523**, 240–244 (2015).
  88. Sabbattini, P. *et al.* A novel role for the Aurora B kinase in epigenetic marking of silent chromatin in differentiated postmitotic cells. *EMBO J.* **26**, 4657–4669 (2007).
  89. Solovei, I. *et al.* LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* **152**, 584–598 (2013).
  90. Rehen, S. K. *et al.* Chromosomal variation in neurons of the developing and adult mammalian nervous system. *Proc. Natl. Acad. Sci. USA* **98**, 13361–13366 (2001).
  91. Rutledge, M. T., Russo, M., Belton, J.-M., Dekker, J. & Broach, J. R. The yeast genome undergoes significant topological reorganization in quiescence. *Nucleic Acids Res.* **43**, 8299–8313 (2015).
  92. Shah, P. P. *et al.* Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. *Genes Dev.* **27**, 1787–1799 (2013).
  93. Sadaie, M. *et al.* Redistribution of the lamin B1 genomic binding profile affects rearrangement of heterochromatic domains and SAHF formation during senescence. *Genes Dev.* **27**, 1800–1808 (2013).
  94. Chandra, T. *et al.* Global reorganization of the nuclear landscape in senescent cells. *Cell Rep.* **10**, 471–483 (2015).
  95. Joshi, O. *et al.* Dynamic reorganization of extremely long-range promoter-promoter interactions between two states of pluripotency. *Cell Stem Cell* **17**, 748–757 (2015).
  96. Lin, Y. C. *et al.* Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. *Nat. Immunol.* **13**, 1196–1204 (2012).
  97. Le Dilys, F. *et al.* Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* **28**, 2151–2162 (2014).
  98. Amano, T. *et al.* Chromosomal dynamics at the *Shh* locus: limb bud-specific differential regulation of competence and active transcription. *Dev. Cell* **16**, 47–57 (2009).
  99. Andrey, G. *et al.* A switch between topological domains underlies *HoxD* genes collinearity in mouse limbs. *Science* **340**, 1234167 (2013).
  100. Montavon, T. *et al.* A regulatory archipelago controls *Hox* genes transcription in digits. *Cell* **147**, 1132–1145 (2011).
  101. Ghavi-Helm, Y. *et al.* Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **512**, 96–100 (2014). **In this paper, 4C is used to interrogate the interaction from approximately 100 enhancers in two developmental stages in *Drosophila melanogaster*. Surprisingly, most interactions seem unchanged and the authors suggest that transcriptional activation is accompanied by a release of paused RNApII from pre-formed enhancer–promoter loops.**
  102. Guelen, L. *et al.* Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948–951 (2008).
  103. Kind, J. *et al.* Single-cell dynamics of genome-nuclear lamina interactions. *Cell* **153**, 178–192 (2013).
  104. Peric-Hupkes, D. *et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* **38**, 603–613 (2010).
  105. Shachar, S., Voss, T. C., Pegoraro, G., Sciascia, N. & Misteli, T. Identification of gene positioning factors using high-throughput imaging mapping. *Cell* **162**, 911–923 (2015).
  106. Deng, W. *et al.* Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233–1244 (2012).
  107. Deng, W. *et al.* Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* **158**, 849–860 (2014). **In this paper, forced chromatin looping between the β-globin gene and its LCR cause upregulation of β-globin transcription, establishing a causal relationship between chromatin looping and gene expression.**
  108. Lupiáñez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
  109. Flavahan, W. A. *et al.* Insulator dysfunction and oncogene activation in *IDH1* mutant gliomas. *Nature* **529**, 110–114 (2016). **In this study, mutations in TAD boundaries are shown to rewire long-range enhancer–promoter interactions and to result in pathogenic phenotypes.**
  110. Hnisz, D. *et al.* Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
  111. Grubert, F. *et al.* Genetic control of chromatin states in humans involves local and distal chromosomal interactions. *Cell* **162**, 1051–1065 (2015).
  112. Waszak, S. M. *et al.* Population variation and genetic control of modular chromatin architecture in humans. *Cell* **162**, 1039–1050 (2015).
  113. Phillips-Cremins, J. E. & Corces, V. G. Chromatin insulators: linking genome organization to cellular function. *Mol. Cell* **50**, 461–474 (2013).
  114. Fanucchi, S., Shibayama, Y., Burd, S., Weinberg, M. S. & Mhlanga, M. M. Chromosomal contact permits transcription between coregulated genes. *Cell* **155**, 606–620 (2013).
  115. Duan, Z. *et al.* A three-dimensional model of the yeast genome. *Nature* **465**, 363–367 (2010).
  116. Le, T. B. K., Imakaev, M. V., Mirny, L. A. & Laub, M. T. High-resolution mapping of the spatial organization of a bacterial chromosome. *Science* **342**, 731–734 (2013).
  117. Mizuguchi, T. *et al.* Cohesin-dependent globules and heterochromatin shape 3D genome architecture in *S. pombe*. *Nature* **516**, 432–435 (2015).
  118. Grob, S., Schmid, M. W. & Grossniklaus, U. Hi-C analysis in *Arabidopsis* identifies the *KNOT*, a structure with similarities to the *flamenco* locus of *Drosophila*. *Mol. Cell* **55**, 678–693 (2014).
  119. Feng, S. *et al.* Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in *Arabidopsis*. *Mol. Cell* **55**, 694–707 (2014).
  120. Véron, A. S., Lemaitre, C., Gautier, C., Lacroix, V. & Sagot, M.-F. Close 3D proximity of evolutionary breakpoints argues for the notion of spatial synteny. *BMC Genomics* **12**, 303 (2011).
  121. Murphy, W. J. *et al.* Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps. *Science* **309**, 613–617 (2005).
  122. Hinsch, H. & Hannenhalli, S. Recurring genomic breaks in independent lineages support genomic fragility. *BMC Evol. Biol.* **6**, 90 (2006).
  123. Gordon, L. *et al.* Comparative analysis of chicken chromosome 28 provides new clues to the evolutionary fragility of gene-rich vertebrate regions. *Genome Res.* **17**, 1603–1613 (2007).
  124. Berthelot, C., Muffato, M., Abecassis, J. & Roest Crollius, H. The 3D organization of chromatin explains evolutionary fragile genomic regions. *Cell Rep.* **10**, 1913–1924 (2015).
  125. Fabre, P. J. *et al.* Nanoscale spatial organization of the *HoxD* gene cluster in distinct transcriptional states. *Proc. Natl. Acad. Sci. USA* **112**, 13964–13969 (2015).
  126. Williamson, I. *et al.* Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence *in situ* hybridization. *Genes Dev.* **28**, 2778–2791 (2014).
  127. Zhu, J. *et al.* Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* **152**, 642–654 (2013).
  128. Wright, A. V., Nuñez, J. K. & Doudna, J. A. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* **164**, 29–44 (2016).
  129. Beliveau, B. J. *et al.* Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl. Acad. Sci. USA* **109**, 21301–21306 (2012).
  130. Beliveau, B. J. *et al.* Single-molecule super-resolution imaging of chromosomes and *in situ* haplotyping visualization using Oligopaint FISH probes. *Nat. Commun.* **6**, 7147 (2015).
  131. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–795 (2006).
  132. Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).
  133. Lööodicé, I., Dubarry, M. & Taddei, A. Scoring and manipulating gene position and dynamics using FROS in budding yeast. *Curr. Protoc. Cell Biol.* **62**, 22.17.1–22.17.14 (2014).
  134. Saad, H. *et al.* DNA dynamics during early double-strand break processing revealed by non-intrusive imaging of living cells. *PLoS Genet.* **10**, e1004187 (2014).
  135. Cullen, K., Kladde, M. & Seyfred, M. Interaction between transcription regulatory regions of prolactin chromatin. *Science* **261**, 203–206 (1993).
  136. Dekker, J. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002).
  137. Fullwood, M. J. *et al.* An oestrogen-receptor-α-bound human chromatin interactome. *Nature* **462**, 58–64 (2009).
  138. Ma, W. *et al.* Fine-scale chromatin interaction maps reveal the *cis*-regulatory landscape of human lincRNA genes. *Nat. Methods* **12**, 71–78 (2015).
  139. van de Werken, H. J. G. *et al.* Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* **9**, 969–972 (2012).
  140. Dostie, J. & Dekker, J. Mapping networks of physical

- interactions between genomic elements using 5C technology. *Nat. Protoc.* **2**, 988–1002 (2007).
141. Hughes, J. R. *et al.* Analysis of hundreds of *cis*-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* **46**, 205–212 (2014).
142. Mifsud, B. *et al.* Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat. Genet.* **47**, 598–606 (2015).
143. Sati, S. & Cavalli, G. Chromosome conformation capture technologies and their impact in understanding genome function. *Chromosoma* <http://dx.doi.org/10.1007/s00412-016-0593-6> (2016).
144. Durand, N. C. *et al.* Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.* **3**, 99–101 (2016).

#### Acknowledgements

Research in the laboratory of G.C. was supported by grants from the European Horizon 2020 (H2020) MuG project under grant agreement No 676556, the Centre National de la Recherche Scientifique (CNRS), the European Network of Excellence EpiGeneSys, the Agence Nationale de la Recherche (EpiDevoMati), the Fondation pour la Recherche Médicale (DE120151234396), the INSERM/Plan Cancer Epigenetics and cancer program (grant acronym "MM&TT"), the Laboratory of Excellence EpiGenMed, and the Fondation ARC pour la Recherche sur le Cancer. B.B. was funded by Sir

Henry Wellcome Postdoctoral Fellowship WT100136MA.

#### Competing interests statement

The authors declare no competing interests.

#### DATABASES

Protein Data Bank: <http://www.rcsb.org/pdb/home/home.do>

#### FURTHER INFORMATION

The Cancer Genome Atlas: <http://cancergenome.nih.gov>

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

## CORRIGENDUM

**Organization and function of the 3D genome***Boyan Bonev and Giacomo Cavalli**Nature Reviews Genetics* 17, 661–678 (2016)

In the original version of this article, the statement that CCCTC-binding factor (CTCF) is conserved in most bilaterians was incorrectly referenced. Reference 58 has now been corrected in the online version of the article to cite Heger, P., Marin, B., Bartkuhn, M., Schierenberg, E. & Wiehe, T. The chromatin insulator CTCF and the emergence of metazoan diversity. *Proc. Natl Acad. Sci. USA* 109, 17507–17512 (2012). The authors apologize for this error.